Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Jan 9 09:33:03 1999; MasPar time 1049.47 Seconds
1495.136 Million cell updates/sec

Tabular output not generated.

Title: >US-08-704-178-1
Description: (1-712) from US08704178.seq (1 of 4)
Perfect Score: 712
N.A. Sequence: 1 ATGGACCTGCAGCTGACCCA.....GGACCACGGTCACCGTCTCC 712
Comp: TACCTGGACGTCGACTGGGT.....CCTGGTGCCAGTGGCAGAGG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 567134 seqs, 1101898692 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl55
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_om 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_vi
Database: genbank107
15:gb_ba1 16:gb_ba2 17:gb_htg 18:gb_in 19:gb_om 20:gb_ov
21:gb_pat 22:gb_ph 23:gb_pl1 24:gb_pl2 25:gb_pr1
26:gb_pr2 27:gb_pr3 28:gb_ro 29:gb_st 30:gb_sts 31:gb_sy
32:gb_un 33:gb_vi

Statistics: Mean 10.607; Variance 5.121; scale 2.071

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	705	99.0	711	21	I32406	Sequence 1 from patent	0.00e+00
2	705	99.0	711	31	S39590	anti-erbB2 imnmunotoxi	0.00e+00
3	410	57.6	720	21	E13599	DNA encoding a signal	0.00e+00
4	375	52.7	720	21	I32407	Sequence 2 from patent	4.82e-291
5	363	51.0	732	21	I31035	Sequence 1 from patent	1.52e-280
6	363	51.0	732	10	E10361	cDNA encoding a monocl	1.52e-280
7	283	39.7	336	28	MMU37860	Mus musculus rearrange	7.59e-211
8	279	39.2	729	10	E10362	cDNA encoding an monoc	2.23e-207
9	279	39.2	729	21	I31036	Sequence 3 from patent	2.23e-207
10	271	38.1	321	28	MMIG013	Mouse hybridoma 2d3 mR	1.91e-200
11	270	37.9	315	21	A36642	Sequence 2 from Patent	1.40e-199
12	270	37.9	315	21	I51657	Sequence 3 from patent	1.40e-199
13	268	37.6	312	28	S48339	Ig V kappa =anti-idiot	7.55e-198
14	268	37.6	318	28	AF056217	Mus musculus monoclona	7.55e-198
15	268	37.6	1370	28	MUSIGKAAR	Mouse Ig kappa active	7.55e-198
16	267	37.5	744	31	SYNDA44X	Synthetic single-chain	5.54e-197
17	266	37.4	354	28	S78745	Ig VH=anti-phosphatidy	4.07e-196
18	265	37.2	345	28	MUSIGKVA2A	Mouse hybridoma Ig rea	2.98e-195
19	265	37.2	351	28	MMU25678	Mus musculus Ig heavy	2.98e-195
20	263	36.9	354	28	MMAHCVR2	M.musculus antibody he	1.60e-193
21	263	36.9	403	28	MUSIGKCNA	Mouse Ig rearranged ka	1.60e-193
22	263	36.9	403	21	I09200	Sequence 40 from Paten	1.60e-193
23	263	36.9	477	28	S76823	Ig V kappa =anti-sigma	1.60e-193
24	262	36.8	321	21	I27490	Sequence 4 from patent	1.17e-192
25	261	36.7	403	21	I05923	Sequence 39 from Paten	8.59e-192
26	260	36.5	355	28	S74055	IgG VH IIA=IgG2b VH re	6.29e-191
27	260	36.5	357	28	MUSCMVG	Mouse mRNA sequence, p	6.29e-191
28	260	36.5	414	21	I78572	Sequence 84 from paten	6.29e-191
29	260	36.5	414	21	I31960	Sequence 84 from paten	6.29e-191
30	260	36.5	414	21	I78627	Sequence 84 from paten	6.29e-191
31	259	36.4	421	28	S51851	M104E myeloma immunogl	4.61e-190
32	258	36.2	309	28	MMIGVKP4	Mouse Ig active L-chai	3.37e-189
33	258	36.2	723	28	AF035617	Mus musculus scFv anti	3.37e-189
34	257	36.1	302	28	MUSIGKAFE	Mus musculus Mouse Ig	2.47e-188
35	257	36.1	403	21	I08812	Sequence 14 from Paten	2.47e-188
36	257	36.1	617	31	XXU05273	Synthetic construct cl	2.47e-188
37	255	35.8	327	28	AF087025	Mus musculus hybrid 2-	1.32e-186
38	255	35.8	384	28	MMU05285	Mus musculus BALB/c/J	1.32e-186
39	255	35.8	732	21	A51863	Sequence 27 from Paten	1.32e-186
40	254	35.7	321	28	MMU60466	Mus musculus Ig anti-D	9.63e-186
41	253	35.5	435	21	A17967	Variable region of the	7.04e-185
42	252	35.4	300	28	MUSIGKAFI	Mus musculus Ig active	5.14e-184
43	252	35.4	366	28	AF081541	Mus musculus anti-HIV-	5.14e-184
44	252	35.4	386	21	A20585	A5B7 humanised mouse m	5.14e-184
45	251	35.3	440	28	MUSHA4	Mouse active H-chain V	3.75e-183

ALIGNMENTS

RESULT 1
LOCUS I32406 711 bp DNA PAT 07-JAN-1997
DEFINITION Sequence 1 from patent US 5587458.
ACCESSION I32406
NID g1823197
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 711)
AUTHORS King,C.Richter, Kasprzyk,P.G. and Bird,R.E.
TITLE Anti-erbB-2 antibodies, combinations thereof, and therapeutic and diagnostic uses thereof
JOURNAL Patent: US 5587458-A 1 24-DEC-1996;
FEATURES Location/Qualifiers
source 1. .711
/organism="unknown"
BASE COUNT 175 a 182 c 190 g 164 t
ORIGIN

Query Match 99.0%; Score 705; DB 21; Length 711;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 710; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

```
Db       1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT- 59
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db      60 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 119
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      61 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 120

Db     120 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTGGCTTCTGGAGTCCCTGC 179
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Qy     121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180

Db     180 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 239
```

```
Qy    181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db    240 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 299

Qy    241 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 300

Db    300 GGGGTCCAAGCTGGAAATAAAAGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAA 359

Qy    301 GGGGTCCAAGCTGGAAATAAAAGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAA 360

Db    360 AGGTGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGAT 419

Qy    361 AGGTGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGAT 420

Db    420 ATCCTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAG 479

Qy    421 ATCCTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAG 480

Db    480 CCATGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTA 539

Qy    481 CCATGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTA 540

Db    540 CAACCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTA 599

Qy    541 CAACCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTA 600

Db    600 CATGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGT 659

Qy    601 CATGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGT 660

Db    660 TACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTCC 711

Qy    661 TACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTCC 712
```

```
RESULT    2
LOCUS       S39590        711 bp                    SYN       10-FEB-1993
DEFINITION  anti-erbB2 imnmunotoxin antigen binding region [mice, Other
            Synthetic Recombinant Partial, 711 nt].
ACCESSION   S39590
NID         g251113
KEYWORDS    .
SOURCE      Mus sp.
  ORGANISM  Mus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 711)
  AUTHORS   Batra,J.K., Kasprzyk,P.G., Bird,R.E., Pastan,I. and King,C.R.
  TITLE     Recombinant anti-erbB2 immunotoxins containing Pseudomonas exotoxin
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 89 (13), 5867-5871 (1992)
  MEDLINE   92335198
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 108547] from the original journal article.
            This sequence comes from Fig. 1.
FEATURES             Location/Qualifiers
     source          1. .711
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
     gene            1. .711
                     /partial
                     /gene="anti-erbB2 imnmunotoxin antigen binding region"
     CDS             1. .711
                     /partial
                     /gene="anti-erbB2 imnmunotoxin antigen binding region"
                     /note="This sequence comes from Fig. 1."
                     /codon_start=1
                     /db_xref="PID:g251114"
                     /translation="MDLQLTQSPAILSASPGEKVTMTCRATPSVSYMHWYQQKPGSSP
                     KPWIYTTSNLASGVPARFSGGGSGTSYSLTVSRVEAEDAATYYCQQWSRSPPTFGGGS
                     KLEIKGSTSGSGKSSEGKGVQLQESGPEVVKPGGSMKISCKTSGYSFTGHTMNWVKQS
                     HGKNLEWIGLINPYNGDTNYNQKFKGKATFTVDKSSSTAYMELLSLTSEDSAVYYCAR
                     RVTDWYFDVWGAGTTVTVS"
BASE COUNT      175 a    182 c    190 g    164 t
ORIGIN
```

```
  Query Match           99.0%;  Score 705;  DB 31;  Length 711;
  Best Local Similarity 99.7%;  Pred. No. 0.00e+00;
  Matches   710;  Conservative    1;  Mismatches   0;  Indels   1;  Gaps   1;

Db      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT- 59

Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db     60 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 119

Qy     61 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 120

Db    120 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTGGCTTCTGGAGTCCCTGC 179

Qy    121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180

Db    180 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 239

Qy    181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db    240 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 299

Qy    241 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 300

Db    300 GGGGTCCAAGCTGGAAATAAAAGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAA 359

Qy    301 GGGGTCCAAGCTGGAAATAAAAGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAA 360

Db    360 AGGTGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGAT 419

Qy    361 AGGTGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGAT 420

Db    420 ATCCTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAG 479

Qy    421 ATCCTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAG 480

Db    480 CCATGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTA 539

Qy    481 CCATGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTA 540

Db    540 CAACCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTA 599

Qy    541 CAACCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTA 600

Db    600 CATGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGT 659

Qy    601 CATGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGT 660

Db    660 TACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTCC 711

Qy    661 TACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTCC 712
```

```
RESULT    3
LOCUS       E13599        720 bp    DNA             PAT       27-APR-1998
DEFINITION  DNA encoding a signal chain Fv antibody of anti-gp130 antibody
            GPX7.
ACCESSION   E13599
NID         g3252401
KEYWORDS    JP 1997220092-A/2.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 720)
  AUTHORS   Ekida,T., Yasukawa,K., Imanaka,T. and Takagi,M. .
  TITLE     PRODUCTION OF SINGLE-STRAND FV ANTIBODY
  JOURNAL   Patent: JP 1997220092-A 2 26-AUG-1997;
            TOSOH CORP
COMMENT     OS   None
```

```
            OC   Artificial sequences.
            PN   JP 1997220092-A/2
            PD   26-AUG-1997
            PF   15-FEB-1996 JP 1996027622
            PI   EKIDA TEIJI, YASUKAWA KIYOSHI, IMANAKA TADAYUKI, PI   TAKAGI
            MASAHIRO
            PC   C12N15/09,C12N1/21,C12P21/08,(C12N1/21,C12R1:19),(C12P21/08,
            PC   C12R1:19);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .720
            FT                   /organism='Antificial sequences'.
FEATURES             Location/Qualifiers
     source          1. .720
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      169 a    193 c    190 g    168 t
ORIGIN

  Query Match           57.6%;  Score 410;  DB 21;  Length 720;
  Best Local Similarity  83.1%;  Pred. No. 0.00e+00;
  Matches   537;  Conservative    1;  Mismatches 104;  Indels   4;  Gaps   3;

Db       9 GCTCACCCAGTCTCCAGCATTGATATCTGCATCTCCAGGGGAGAAGGT-CACCATGACCT 67
          ||| ||||||||||||||| |  | |||||||||||||||||||| ||| ||||| |
Qy      12 GCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTT 71

Db      68 GCAATGTCAGCTCAAGTGTTACTTCCATGTATTGGTACCAGCAGAAGCCAGGATCCTCCC 127
          |||  | || | ||||||| | || |||| | ||||| ||||||||||||||||||||||
Qy      72 GCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCC 131

Db     128 CCAAACCCTGGATTTATCTCACATCCAACCTGGCTTCTGGAGTCCCTCCTCGCTCCAGTG 187
          ||||||| |||||||||  ||||||||||||:|||||||||||||||| ||||||  |||||
Qy     132 CCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTG 191

Db     188 GCAGTGGGTCTGGGACCACTTACTCTCTCACAATCAGCAGCTTGGAGGCTGAAGATGCTG 247
          || ||||||||||||| ||||||||||||||| |||||||  ||||||||||||||||||
Qy     192 GCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTG 251

Db     248 CCACTTATTACTGCCAGCAGTGGAGTACTAACCCGCTCACGTTCGGTGCTGGGACCAAGC 307
          |||||||||||||||||||||||||||  || ||| | |||||||| |  ||| ||||||
Qy     252 CCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGC 311

Db     308 TCGAGCTGGGTGGCGGTGGCTCGGGCGGTGGTGGGTCGGGTGGCGGCGGATCGGACGTCC 367
          | ||  |    ||   |  ||| ||   ||||   ||   ||  ||   |  ||  || |
Qy     312 TGGAAATAAAAGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAA--GGT-GTGC 368

Db     368 AATTGCAGCAGTCTGGACCTGAACTGGTGAAGCCTGGGGCTTCAGTGAAGATACCCTGCA 427
          |  ||||| |||| |||||||| ||||||||||||| | |||| |||||||| ||||||
Qy     369 AGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCCTGCA 428

Db     428 AGGCTTCAGGATACACATTCACTGACTACAACATGGACTGGGTGAAGCAGAGCCATGGAA 487
          || |||| || ||| |||||||||  | ||| |||| ||||||||||||||||||||||||
Qy     429 AGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCATGGAA 488

Db     488 AGAGCCTTGAGTGGATTGGAGATATTAATTCTCATAGTGGTGGTATTATCTACAACCAAA 547
          ||| |||||||||||||||  ||||||| || | | |||||  || || |||||||||  |
Qy     489 AGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAACCAGA 548

Db     548 AGTTCAAGGACAAGGCCATATTGACTGTAGATAAGTCCTCCAGCACAGCCTACATGGAGC 607
          |||||||| |||||||| ||| |||||||| ||||| ||||||||||||||||||||||||
Qy     549 AGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATGGAGC 608

Db     608 TCCGCAGCCTGACATCTGAGGACACTGCAGTCTATTATTGTGCAAG 653
          ||| ||| ||||||||||||||| |||||||||||| ||||||||
Qy     609 TCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654


RESULT    4
```

```
LOCUS       I32407        720 bp    DNA             PAT       07-JAN-1997
DEFINITION  Sequence 2 from patent US 5587458.
ACCESSION   I32407
NID         g1823198
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 720)
  AUTHORS   King,C.Richter, Kasprzyk,P.G. and Bird,R.E.
  TITLE     Anti-erbB-2 antibodies, combinations thereof, and therapeutic and
            diagnostic uses thereof
  JOURNAL   Patent: US 5587458-A 2 24-DEC-1996;
FEATURES             Location/Qualifiers
     source          1. .720
                     /organism="unknown"
BASE COUNT      184 a    178 c    179 g    179 t
ORIGIN

  Query Match           52.7%;  Score 375;  DB 21;  Length 720;
  Best Local Similarity  79.3%;  Pred. No. 4.82e-291;
  Matches   511;  Conservative    1;  Mismatches 131;  Indels   1;  Gaps   1;

Db       2 TGCAGCTGACCCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAAAAGGT-CACCATG 60
          |||||||||||||||||||||||||||||-|||||||||||||||||| ||||| ||| |||
Qy       8 TGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATG 67

Db      61 ACCTGCAGTGCCAGCTCAAGTGTAAGTAACATGCACTGGTATCAGCAGAAGTCAAGCACC 120
          || ||||| |||| | |||||||||||| |||||||||||||||||||||| || |  ||
Qy      68 ACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCC 127

Db     121 TCCCCCAAACTCTGGGTTTATGACACATCCAAACTGGCTTCTGGAGTCCCAGGTCGCTTC 180
          |||||||||  ||| |||||  ||||||||| ||:||||||||||||| | |||||||
Qy     128 TCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTC 187

Db     181 AGTGGCAGTGGGTCTGGAAACTCTTACTCTCTCACGATCAGCAGCATGGAGGCTGAAGAT 240
          |||||| |||||||||| | ||||||||||||||  |||||||  ||||||||||||||
Qy     188 AGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGAT 247

Db     241 GCTGCCACTTATTATTGTTATCAGGGGAGTGGGTACCCATTCACGTTCGGCTCGGGGACA 300
          |||||||||||||| ||  | ||| ||||| |   ||||  |||||||||   |||| |
Qy     248 GCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCC 307

Db     301 AAGTTGGAAATAAAAGGTTCTACCTCCGGATCTGGTAAATCTTCTGAAGGTAAAGGTGTG 360
          ||| ||||||||||||||||||||| || ||||||||||||||||||||||||||||||||
Qy     308 AAGCTGGAAATAAAAGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGGTGTG 367

Db     361 CAGCTGCAGCAGTCTGGGGTTGAGCTTGTCCGAGGAGGGGCCTTAGTCAAGTTGTCCTGC 420
          ||||||||| |||| ||   |||| | ||       || |  | | | ||| | ||||||
Qy     368 CAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCCTGC 427

Db     421 AAAGCTTCTGACTTCAACATTAAAGACTATTATATCCACTGGGTGAAGCAGAGGCCTGAA 480
          ||  ||||||  | |    | |  | | |    ||  |||||||||||||||| | || |
Qy     428 AAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCATGGA 487

Db     481 CAGGGCCTGGAATGGATTGGATGGATTCATCCTGAGAATGGTAATACTGTATATGACCCG 540
          ||  ||| || ||||||||   ||| ||||| | |||||| |||||   ||  ||| |
Qy     488 AAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAACCAG 547

Db     541 AAATTCCAGGGCAAGGCCAGTATAACAGCAGACACATCCTCCAACGCGGCCTACCTTCAG 600
          || ||| ||||||||||||   | || | |||||  || |||| | | ||||||  |  ||
Qy     548 AAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATGGAG 607

Db     601 CTCAGCAGCCTGACATCTGAGGACACTGCCGTCTATTACTGTGC 644
          |||  ||| ||||||||||||||| |||| ||||||||||||||
Qy     608 CTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGC 651


RESULT    5
LOCUS       I31035        732 bp    DNA             PAT       20-DEC-1996
DEFINITION  Sequence 1 from patent US 5582826.
ACCESSION   I31035
NID         g1821826
```

```
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 732)
  AUTHORS   Shimamura,T., Hamuro,J., Nakazawa,H., Kanayama,Y., Sugamura,K. and
            Takeshita,T.
  TITLE     Monoclonal antibodies which bind the gamma chain of human
            interleukin-2 receptor
  JOURNAL   Patent: US 5582826-A 1 10-DEC-1996;
FEATURES             Location/Qualifiers
     source          1. .732
                     /organism="unknown"
BASE COUNT      180 a    204 c    182 g    166 t
ORIGIN

  Query Match           51.0%;  Score 363;  DB 21;  Length 732;
  Best Local Similarity  79.8%;  Pred. No. 1.52e-280;
  Matches   550;  Conservative    1;  Mismatches 128;  Indels  10;  Gaps   4;

Db      1 ATGGACATCCTGCTGACCCAGTCTCCATCAATCATGTCTGCATCTCTAGGGGAACGGGT- 59
          |||||| | | |||||||||||||||| ||||| |||||||||||| ||||||   ||| 
Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db     60 CACCATGACCTGCACTGCCAGCTCAAGTGTAAGTTCCAGTTACTTGCACTGGTACCAGCA 119
          ||| ||||| |||| |||| | |||||||||||| ||      |||||||||| |||||
Qy     61 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACA------TGCACTGGTATCAGCA 114

Db    120 GAAGCCAGGATCCTCCCCCAAACTCTGGATTTATAGCACATCCAACCTGGCTTCTGGAGT 179
          ||||||||||||||||||||  |||||||||| ||||||||||:|||||||||||
Qy    115 GAAGCCAGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGT 174

Db    180 CCCAGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGCAT 239
          ||| |||||||||||||| ||||||||||||||||||||||||||||| |||||||  | 
Qy    175 CCCTGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGT 234

Db    240 GGAGGCTGAAGATGCTGCCACTTATTACTGCCACCAGTATCATCGTTCCCCGCTCACGTT 299
          |||||||||||||||||||||||||||||||| ||||    ||||  ||| | ||||||
Qy    235 GGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTT 294

Db    300 CGGTGCTGGGACCAAGCTGGAGCTCAAAGTCGACAAATCCTCAGGATCTGGCTCCGAATC 359
          ||| |  ||| |||||||||| | ||||    | |  |||| || |||||       ||
Qy    295 CGGAGGGGGGTCCAAGCTGGAAATAAAAGGTT-CTA--CCTCTGGTTCTGGTAAATCTTC 351

Db    360 CAAAAGCACGCAGGTCAAACTCGAGGAGTCTGGATCTGAGCTGGTGAGGCCTGGAGCTTC 419
            || | |     ||  | ||  ||||||| ||| ||||| |||||| ||||||| |||
Qy    352 TGAAGGTAAAGGTGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTC 411

Db    420 AGTGAAGCTGTCCTGCAAGGCTTCTGGCTACACATTCACCAGCTACTGGATGCACTGGGT 479
          | ||||| | |||||||| ||||||| ||| |||||||  || ||   ||| |||||||
Qy    412 AATGAAGATATCCTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGT 471

Db    480 GAAGCAGAGGCATGGACAAGGCCTTGAGTGGATTGGAAATATTTATCCTGGTAGTGGTAG 539
          |||||||| ||||||  |   ||||||||||||||||  |||| |||||    | ||||
Qy    472 GAAGCAGAGCCATGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGA 531

Db    540 TACTAACTACGATGAGAAGTTCAAGAGCAAGGCCACACTGACTGTAGACACATCCTCCAG 599
          |||||||||| |  |||||||||| |||||||||| | |||||||||  || |||||
Qy    532 TACTAACTACAACCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAG 591

Db    600 CACAGCCTACATGCACCTCAGCAGCCTGACATCTGAGGACTCTGCGGTCTATTACTGTAC 659
          |||||||||||| | |||  ||| |||||||||||||||||| ||||||||||| |
Qy    592 CACAGCCTACATGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGC 651

Db    660 AAGAAGCAGCCGGAACTGGGTCTACTATG 688
          ||| ||      | |||||  || | |||
Qy    652 AAGGAGGGTTACGGACTGGTACTTCGATG 680


RESULT    6
ID   E10361     standard; RNA; ROD; 732 BP.
AC   E10361;
NI   d1108698
DT   08-OCT-1997 (Rel. 52, Created)
DT   08-OCT-1997 (Rel. 52, Last updated, Version 1)
DE   cDNA encoding a monoclonal antibody against human Interleukin 2
DE   receptor gamma chain.
KW   JP 1995313188-A/1.
OS   Mus sp.
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia;
OC   Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   1-732
RA   Shimamura T., Hamuro J., Nakazawa H., Kanayama Y., Sugamura K.,
RA   Takeshita T.;
RT   "IMMUNOSUPPRESSIVE AGENT";
RL   Patent number JP 1995313188-A/1, 05-DEC-1995.
RL   AJINOMOTO CO INC, SUGAMURA KAZUO,.
CC   OS   Mus sp. (mouse)
CC   PN   JP 1995313188-A/1
CC   PD   05-DEC-1995
CC   PF   21-APR-1994 JP 1994082836
CC   PR   21-APR-1993 JP 93P    94491, 07-MAR-1994 JP 94P    36065
CC   PI   SHIMAMURA TOSHIAKI, HAMURO JUNJI, NAKAZAWA HARUMI,
CC   PI   KANAYAMA YUKA,
CC   PI   SUGAMURA KAZUO, TAKESHITA TOSHIICHI
CC   PC   C12P21/08,A61K39/395,A61K39/395,C12N1/21,C12N5/20,
CC   PC   C12N15/13//C12N15/06,
CC   PC   (C12P21/08,C12R1:19),(C12P21/08,C12R1:91),(C12N1/21,C12R1:19);
CC   CC   strandedness: Double;
CC   CC   topology: Linear;
CC   FH   Key             Location/Qualifiers
CC   FH
CC   FT   source          1. .732
CC   FT                   /organism="Mus sp."
CC   FT                   /cell_type="hybridoma"
CC   FT                   /cell_line="GP-2"
CC   FT   mat_peptide     1. .732
CC   FT                   /product="anti-IL-2 receptor gamma chain"
FH   Key             Location/Qualifiers
FH
FT   source          1. .732
FT                   /organism="Mus sp."
SQ   Sequence 732 BP; 180 A; 204 C; 182 G; 166 T; 0 other;

  Query Match           51.0%;  Score 363;  DB 10;  Length 732;
  Best Local Similarity  79.8%;  Pred. No. 1.52e-280;
  Matches   550;  Conservative    1;  Mismatches 128;  Indels  10;  Gaps   4;

Db      1 ATGGACATCCTGCTGACCCAGTCTCCATCAATCATGTCTGCATCTCTAGGGGAACGGGT- 59
          |||||| | | |||||||||||||||| ||||| |||||||||||| ||||||   ||| 
Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db     60 CACCATGACCTGCACTGCCAGCTCAAGTGTAAGTTCCAGTTACTTGCACTGGTACCAGCA 119
          ||| ||||| |||| |||| | |||||||||||| ||      |||||||||| |||||
Qy     61 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACA------TGCACTGGTATCAGCA 114

Db    120 GAAGCCAGGATCCTCCCCCAAACTCTGGATTTATAGCACATCCAACCTGGCTTCTGGAGT 179
          ||||||||||||||||||||  |||||||||| ||||||||||:|||||||||||
Qy    115 GAAGCCAGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGT 174

Db    180 CCCAGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGCAT 239
          ||| |||||||||||||| ||||||||||||||||||||||||||||| |||||||  | 
Qy    175 CCCTGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGT 234

Db    240 GGAGGCTGAAGATGCTGCCACTTATTACTGCCACCAGTATCATCGTTCCCCGCTCACGTT 299
          |||||||||||||||||||||||||||||||| ||||    ||||  ||| | ||||||
Qy    235 GGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTT 294

Db    300 CGGTGCTGGGACCAAGCTGGAGCTCAAAGTCGACAAATCCTCAGGATCTGGCTCCGAATC 359
          ||| |  ||| |||||||||| | ||||    | |  |||| || |||||       ||
Qy    295 CGGAGGGGGGTCCAAGCTGGAAATAAAAGGTT-CTA--CCTCTGGTTCTGGTAAATCTTC 351

Db    360 CAAAAGCACGCAGGTCAAACTCGAGGAGTCTGGATCTGAGCTGGTGAGGCCTGGAGCTTC 419
            || | |     ||  | ||  ||||||| ||| ||||| |||||| ||||||| |||
Qy    352 TGAAGGTAAAGGTGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTC 411
```

```
Db    420 AGTGAAGCTGTCCTGCAAGGCTTCTGGCTACACATTCACCAGCTACTGGATGCACTGGGT 479
          | ||||| | |||||||||| ||||||| ||| |||||||  || ||   ||| |||||||
Qy    412 AATGAAGATATCCTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGT 471

Db    480 GAAGCAGAGGCATGGACAAGGCCTTGAGTGGATTGGAAATATTTATCCTGGTAGTGGTAG 539
          ||||||||| |||||| |   ||||||||||||||||  |||| |||||   | ||||
Qy    472 GAAGCAGAGCCATGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGA 531

Db    540 TACTAACTACGATGAGAAGTTCAAGAGCAAGGCCACACTGACTGTAGACACATCCTCCAG 599
          ||||||||||  |  ||||||||||| |||||||||| | |||||||||   || |||||
Qy    532 TACTAACTACAACCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAG 591

Db    600 CACAGCCTACATGCACCTCAGCAGCCTGACATCTGAGGACTCTGCGGTCTATTACTGTAC 659
          |||||||||||||  | |||  ||| |||||||||||||||||||| ||||||||||| |
Qy    592 CACAGCCTACATGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGC 651

Db    660 AAGAAGCAGCCGGAACTGGGTCTACTATG 688
          ||| ||      | |||||  || | |||
Qy    652 AAGGAGGGTTACGGACTGGTACTTCGATG 680


RESULT    7
LOCUS       MMU37860      336 bp    mRNA            ROD       16-MAR-1996
DEFINITION  Mus musculus rearranged immunoglobulin heavy chain variable region
            mRNA, partial cds.
ACCESSION   U37860
NID         g1127634
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 336)
  AUTHORS   Caton,A.J., Swartzentruber,J.R., Kuhl,A.L., Carding,S.R. and
            Stark,S.E.
  TITLE     Activation and negative selection of functionally distinct subsets
            of antibody-secreting cells by influenza hemagglutinin as a viral
            and a neo-self antigen
  JOURNAL   J. Exp. Med. 183 (1), 13-26 (1996)
  MEDLINE   96136744
REFERENCE   2  (bases 1 to 336)
  AUTHORS   Swartzentruber,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-OCT-1995) Jennifer R. Swartzentruber, The Wistar
            Institute, 3601 Spruce St., Philadelphia, PA 19104, USA
FEATURES             Location/Qualifiers
     source          1. .336
                     /organism="Mus musculus"
                     /isolate="h922-163"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
                     /cell_type="B-lymphocyte"
     CDS             <1. .>336
                     /codon_start=1
                     /product="rearranged immunoglobulin heavy chain variable
                     region"
                     /db_xref="PID:g1127635"
                     /translation="ELVKPGASMKISCKASGYSFTGYTMNWVKQSHGKNLEWIGLINP
                     YNGGTSYNQKFKGKATLTVDKSSSTAYMELLSLTSEDSAVYYCAREWLLRYFDVWGAG
                     TTVTVSSAKT"
BASE COUNT       92 a     81 c     85 g     78 t
ORIGIN

  Query Match           39.7%;  Score 283;  DB 28;  Length 336;
  Best Local Similarity  95.1%;  Pred. No. 7.59e-211;
  Matches   309;  Conservative    0;  Mismatches  14;  Indels   2;  Gaps   2;

Db      1 GAGCTGGTGAAGCCTGGAGCTTCAATGAAGATATCCTGCAAGGCTTCTGGTTACTCATTC 60
          ||| ||||||||||||||| ||||||||||||||||||||| ||||||||||||||||||
Qy    389 GAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCCTGCAAGACTTCTGGTTACTCATTC 448

Db     61 ACTGGCTACACCATGAACTGGGTGAAGCAGAGCCATGGAAAGAACCTTGAGTGGATTGGA 120
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    449 ACTGGCCACACCATGAACTGGGTGAAGCAGAGCCATGGAAAGAACCTTGAGTGGATTGGA 508

Db    121 CTTATTAATCCTTACAATGGTGGTACTAGCTACAACCAGAAGTTCAAGGGCAAGGCCACA 180
          |||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||
Qy    509 CTTATTAATCCTTACAATGGTGATACTAACTACAACCAGAAGTTCAAGGGCAAGGCCACA 568

Db    181 TTAACTGTAGACAAGTCATCCAGCACAGCCTACATGGAGCTCCTCAGTCTGACATCTGAG 240
          || ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Qy    569 TTTACTGTAGACAAGTCGTCCAGCACAGCCTACATGGAGCTCCTCAGTCTGACATCTGAG 628

Db    241 GACTCTGCAGTCTATTACTGTGCAAGAGAGTGGTTACTAC-GGTACTTCGATGTCTGGGG 299
          |||||||||||||||||||||||||| ||| | |    || ||||||||||||||||||
Qy    629 GACTCTGCAGTCTATTACTGTGCAAG-GAGGGTTACGGACTGGTACTTCGATGTCTGGGG 687

Db    300 CGCAGGGACCACGGTCACCGTCTCC 324
          |||||||||||||||||||||||||
Qy    688 CGCAGGGACCACGGTCACCGTCTCC 712


RESULT    8
ID   E10362     standard; RNA; ROD; 729 BP.
AC   E10362;
NI   d1108699
DT   08-OCT-1997 (Rel. 52, Created)
DT   08-OCT-1997 (Rel. 52, Last updated, Version 1)
DE   cDNA encoding an monoclonal antibody against human Interleukin-2
DE   receptor gamma chain.
KW   JP 1995313188-A/2.
OS   Mus sp.
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia;
OC   Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   1-729
RA   Shimamura T., Hamuro J., Nakazawa H., Kanayama Y., Sugamura K.,
RA   Takeshita T.;
RT   "IMMUNOSUPPRESSIVE AGENT";
RL   Patent number JP 1995313188-A/2, 05-DEC-1995.
RL   AJINOMOTO CO INC, SUGAMURA KAZUO,.
CC   OS   Mus sp. (mouse)
CC   PN   JP 1995313188-A/2
CC   PD   05-DEC-1995
CC   PF   21-APR-1994 JP 1994082836
CC   PR   21-APR-1993 JP 93P    94491, 07-MAR-1994 JP 94P    36065
CC   PI   SHIMAMURA TOSHIAKI, HAMURO JUNJI, NAKAZAWA HARUMI,
CC   PI   KANAYAMA YUKA,
CC   PI   SUGAMURA KAZUO, TAKESHITA TOSHIICHI
CC   PC   C12P21/08,A61K39/395,A61K39/395,C12N1/21,C12N5/20,
CC   PC   C12N15/13//C12N15/06,
CC   PC   (C12P21/08,C12R1:19),(C12P21/08,C12R1:91),(C12N1/21,C12R1:19);
CC   FH   Key             Location/Qualifiers
CC   FH
CC   FT   source          1. .729
CC   FT                   /organism="Mus sp."
CC   FT                   /cell_type="hybridoma"
CC   FT                   /cell_line="GP-4"
CC   FT   mat_peptide     1. .729
CC   FT                   /product="anti-IL-2 receptor gamma chain"
FH   Key             Location/Qualifiers
FH
FT   source          1. .729
FT                   /organism="Mus sp."
SQ   Sequence 729 BP; 199 A; 169 C; 182 G; 179 T; 0 other;

  Query Match           39.2%;  Score 279;  DB 10;  Length 729;
  Best Local Similarity  73.8%;  Pred. No. 2.23e-207;
  Matches   487;  Conservative    0;  Mismatches 166;  Indels   7;  Gaps   5;

Db      1 ATGGATATTCTGCTGACACAGTCTCCAGCCTCCCTATCTGCATCTGTGGGAGAAACTGT- 59
          |||||  | | |||||| ||||||||||   ||| |||||||||   || || |  ||
Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60
```

```
Db      60 CACCATCACATGTCGAGCAAGTGGGAATATTCACAATTATTTAGCATGGTATCAGCAGAA 119
           ||| || || ||  | || |     | | |  | | |||  |    ||||||||||||||
Qy      61 CACAATGACTTGCAGGGCCACCCCAAGTGT--A-AGTTACATGCACTGGTATCAGCAGAA 117

Db     120 ACAGGGAAAATCTCCTCAGCTCCTGGTCTATAATGCAAAAACCTTAGCAGATGGTGTGCC 179
           |  |||   || ||  | |    | | ||||   ||   | | ||   ||| || ||
Qy     118 GCCAGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCC 177

Db     180 ATCAAGGTTCAGTGGCAGTGGATCAGGAACACAATATTCTCTCAAGATCAACAGCCTGCA 239
             |  | ||||||||| |||| || || ||    || |||||||   ||| |||  || |
Qy     178 TGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGA 237

Db     240 GCCTGAAGATTTTGGGAGTTATTACTGTCAACATTTTTGGAGTACTCCGTGGACGTTCGG 299
           | |||||||||   ||  | |||||||||| || || |   |  |||  ||    ||||||||
Qy     238 GGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGG 297

Db     300 TGGAGGGACCAAGCTGGAGCTCAAAGTCGAGAAATCCTCAGGATCTGGCTCCGAATCCAA 359
           || ||| ||||||||||  | ||||  |    | |||| || |||||        ||  |
Qy     298 AGGGGGGTCCAAGCTGGAAATAAAAG--GTTCTA-CCTCTGGTTCTGGTAAATCTTCTGA 354

Db     360 AAGCACGCAGGTCAAACTCGAGGAGTCTGGACCTGAGCTGGTGAAGCCTGGGGCTTCAGT 419
           | | |     ||  | ||  ||||||| |||||||| |||||||||||| | |||| |
Qy     355 AGGTAAAGGTGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAAT 414

Db     420 GAAGATATCCTGCAAGGCTTCTGGTTACTCATTCACTGGCTACTACATGCACTGGGTGAA 479
           |||||||||||||||| |||||||||||||||||||||| ||  |||| ||||||||||
Qy     415 GAAGATATCCTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAA 474

Db     480 GCAAAGCCATGTAAAGAGCCTTGAGTGGATTGGACGTATTAATCCTTACAATGGTGCTAC 539
           ||| ||||||| ||||| ||||||||||||||||| |||||||||||||||||||||| |||
Qy     475 GCAGAGCCATGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATAC 534

Db     540 TAGCTACAACCAGAATTTCAAGGACAAGGCCAGCTTGACTGTAGATAAGTCCTCCAGCAC 599
           || |||||||||||| ||||||| ||||||||  || |||||||| ||||| ||||||||
Qy     535 TAACTACAACCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCAC 594

Db     600 AGCCTACATGGAGCTCCACAGCCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 659
           |||||||||||||||| ||| ||||||||||||||||||||||||||||||||||||||
Qy     595 AGCCTACATGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654


RESULT    9
LOCUS       I31036        729 bp    DNA             PAT       20-DEC-1996
DEFINITION  Sequence 3 from patent US 5582826.
ACCESSION   I31036
NID         g1821827
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 729)
  AUTHORS   Shimamura,T., Hamuro,J., Nakazawa,H., Kanayama,Y., Sugamura,K. and
            Takeshita,T.
  TITLE     Monoclonal antibodies which bind the gamma chain of human
            interleukin-2 receptor
  JOURNAL   Patent: US 5582826-A 3 10-DEC-1996;
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="unknown"
BASE COUNT      199 a    169 c    182 g    179 t
ORIGIN

  Query Match           39.2%;  Score 279;  DB 21;  Length 729;
  Best Local Similarity  73.8%;  Pred. No. 2.23e-207;
  Matches   487;  Conservative    0;  Mismatches 166;  Indels   7;  Gaps   5;

Db       1 ATGGATATTCTGCTGACACAGTCTCCAGCCTCCCTATCTGCATCTGTGGGAGAAACTGT- 59
           |||||  | | |||||| ||||||||||   ||| |||||||||   || || |  ||
Qy       1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db      60 CACCATCACATGTCGAGCAAGTGGGAATATTCACAATTATTTAGCATGGTATCAGCAGAA 119
           ||| || || ||  | || |     | | |  | | |||  |    ||||||||||||||
Qy      61 CACAATGACTTGCAGGGCCACCCCAAGTGT--A-AGTTACATGCACTGGTATCAGCAGAA 117

Db     120 ACAGGGAAAATCTCCTCAGCTCCTGGTCTATAATGCAAAAACCTTAGCAGATGGTGTGCC 179
           |  |||   || ||  | |    | | ||||   ||   | | ||   ||| || ||
Qy     118 GCCAGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCC 177

Db     180 ATCAAGGTTCAGTGGCAGTGGATCAGGAACACAATATTCTCTCAAGATCAACAGCCTGCA 239
             |  | ||||||||| |||| || || ||    || |||||||   ||| |||  || |
Qy     178 TGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGA 237

Db     240 GCCTGAAGATTTTGGGAGTTATTACTGTCAACATTTTTGGAGTACTCCGTGGACGTTCGG 299
           | |||||||||  ||  | |||||||||| || || |   |  |||  ||    ||||||||
Qy     238 GGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGG 297

Db     300 TGGAGGGACCAAGCTGGAGCTCAAAGTCGAGAAATCCTCAGGATCTGGCTCCGAATCCAA 359
           || ||| ||||||||||  | ||||  |    | |||| || |||||        ||  |
Qy     298 AGGGGGGTCCAAGCTGGAAATAAAAG--GTTCTA-CCTCTGGTTCTGGTAAATCTTCTGA 354

Db     360 AAGCACGCAGGTCAAACTCGAGGAGTCTGGACCTGAGCTGGTGAAGCCTGGGGCTTCAGT 419
           | | |     ||  | ||  ||||||| |||||||| |||||||||||| | |||| |
Qy     355 AGGTAAAGGTGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAAT 414

Db     420 GAAGATATCCTGCAAGGCTTCTGGTTACTCATTCACTGGCTACTACATGCACTGGGTGAA 479
           |||||||||||||||| |||||||||||||||||||||| ||  |||| ||||||||||
Qy     415 GAAGATATCCTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAA 474

Db     480 GCAAAGCCATGTAAAGAGCCTTGAGTGGATTGGACGTATTAATCCTTACAATGGTGCTAC 539
           ||| ||||||| ||||| ||||||||||||||||| |||||||||||||||||||||| |||
Qy     475 GCAGAGCCATGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATAC 534

Db     540 TAGCTACAACCAGAATTTCAAGGACAAGGCCAGCTTGACTGTAGATAAGTCCTCCAGCAC 599
           || |||||||||||| ||||||| ||||||||  || |||||||| ||||| ||||||||
Qy     535 TAACTACAACCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCAC 594

Db     600 AGCCTACATGGAGCTCCACAGCCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 659
           |||||||||||||||| ||| ||||||||||||||||||||||||||||||||||||||
Qy     595 AGCCTACATGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654


RESULT   10
LOCUS       MMIG013       321 bp    RNA             ROD       06-AUG-1995
DEFINITION  Mouse hybridoma 2d3 mRNA for immunoglobuin kappa light chain V
            region.
ACCESSION   X58586 Y00794
NID         g51562
KEYWORDS    Ig kappa light chain; Ig variable region; immunoglobulin.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 321)
  AUTHORS   Meek,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-1989) K. Meek, University of Texas Southwestern
            Medical School, 5323 Harry Hines Blud. Microbiology, Dallas Texas
            75235, USA
REFERENCE   2  (bases 1 to 321)
  AUTHORS   Meek,K., Hasemann,C., Pollok,B., Alkan,S.S., Brait,M., Slaoui,M.,
            Urbain,J. and Capra,J.D.
  TITLE     Structural characterization of antiidiotypic antibodies. Evidence
            that Ab2s are derived from the germline differently than Abls
  JOURNAL   J. Exp. Med. 169 (2), 519-533 (1989)
  MEDLINE   89094248
FEATURES             Location/Qualifiers
     source          1. .321
                     /organism="Mus musculus"
                     /strain="Balb/c"
                     /db_xref="taxon:10090"
                     /cell_line="Hybridoma 2D3-K"
     V_region        <1. .>321
                     /gene="IG kappa light chain"
                     /note="variable region"
                     /product="IG kappa light chain"
```

```
     mRNA            <1. .>321
                     /gene="IG kappa light chain"
                     /note="variable region"
                     /evidence=experimental
     gene            1. .321
                     /gene="IG kappa light chain"
     CDS             <1. .>321
                     /gene="IG kappa light chain"
                     /note="variable region"
                     /codon_start=1
                     /product="IG kappa light chain"
                     /db_xref="PID:g938245"
                     /translation="QIVLSQSPAILSASPGEKVTMTCRASSSVSYMHWYQQKPGSSPK
                     PWISATSNLASGVPARFSGSGSGTSYSLTISRVEAEDAATYYCHQWSSNPPTFGGGTK
                     LEIKR"
BASE COUNT       78 a     92 c     80 g     71 t
ORIGIN

  Query Match           38.1%;  Score 271;  DB 28;  Length 321;
  Best Local Similarity  94.8%;  Pred. No. 1.91e-200;
  Matches   290;  Conservative    1;  Mismatches  14;  Indels   1;  Gaps   1;

Db      14 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CACAATGACTTGCAGG 72
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Qy      17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db      73 GCCAGCTCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCAAAA 132
           |||| | ||||||||||||||||||||||||||||||||||||||||||||||||| |||
Qy      77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db     133 CCCTGGATTTCTGCCACATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGT 192
           || ||||||| | |||||||||||||:||||||||||||||||||||||||||||||| ||
Qy     137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db     193 GGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 252
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Qy     197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db     253 TATTACTGCCACCAGTGGAGTAGTAACCCCCCAACGTTCGGAGGGGGGACCAAGCTGGAA 312
           |||||||||||| ||||||||| ||| ||| || ||||||||||||||| ||||||||||
Qy     257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAA 316

Db     313 ATAAAA 318
           ||||||
Qy     317 ATAAAA 322


RESULT   11
LOCUS       A36642        315 bp    DNA             PAT       05-MAR-1997
DEFINITION  Sequence 2 from Patent EP0585570.
ACCESSION   A36642
NID         g2293945
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 315)
  AUTHORS   Seemann,G. and Bosslet,K.
  TITLE     Antigranulocyte antibody construct, preparation and use thereof
  JOURNAL   Patent: EP 0585570-A 2 09-MAR-1994;
            BEHRINGWERKE AG (DE)
COMMENT     Other publication JP 6205693 940726
            Other publication CA 2101868 940206
            Other publication AU 4441393 940210
            Other publication AU 666029 960125
            Other publication DE 4225853 940210.
FEATURES             Location/Qualifiers
     source          1. .315
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /cell_type="hybridoma/myeloma"
BASE COUNT       72 a     92 c     80 g     71 t
ORIGIN

  Query Match           37.9%;  Score 270;  DB 21;  Length 315;
  Best Local Similarity  93.3%;  Pred. No. 1.40e-199;
  Matches   294;  Conservative    1;  Mismatches  19;  Indels   1;  Gaps   1;

Db       1 GACATTCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CAC 59
           ||| | ||||||||||||||||||||||||||||||||||||||||||||||||| |||
Qy       4 GACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACAC 63

Db      60 AATGACTTGCAGGGCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGG 119
           |||||||||||||||| | |||||||||||||||||||||||| ||||||||||||||
Qy      64 AATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGG 123

Db     120 ATCCTCCCCCAAACCCTGGATTTATGCCACATCCAACCTGGCTTCTGGAGTCCCTGCTCG 179
           |||||||||||||| |||||||| ||||||||||||:||||||||||||||||||||||
Qy     124 ATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCG 183

Db     180 CTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCATCAGAGTGGAGGCTGA 239
           |||||||||| |||||||||||||||||||||||||||| ||| ||||||||||||||
Qy     184 CTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGA 243

Db     240 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTAGTAACCCGCTCACGTTCGGTGCTGG 299
           |||||||||||||||||||||||||||||||| ||| ||| | ||||||||| |  ||
Qy     244 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGG 303

Db     300 GACCAAGCTGGAGAT 314
           | |||||||||| ||
Qy     304 GTCCAAGCTGGAAAT 318


RESULT   12
LOCUS       I51657        315 bp    DNA             PAT       12-AUG-1997
DEFINITION  Sequence 3 from patent US 5645817.
ACCESSION   I51657
NID         g2472858
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 315)
  AUTHORS   Seemann,G. and Bosslet,K.
  TITLE     Granulocyte-binding antibody constructs, their preparation and use
  JOURNAL   Patent: US 5645817-A 3 08-JUL-1997;
FEATURES             Location/Qualifiers
     source          1. .315
                     /organism="unknown"
BASE COUNT       72 a     92 c     80 g     71 t
ORIGIN

  Query Match           37.9%;  Score 270;  DB 21;  Length 315;
  Best Local Similarity  93.3%;  Pred. No. 1.40e-199;
  Matches   294;  Conservative    1;  Mismatches  19;  Indels   1;  Gaps   1;

Db       1 GACATTCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CAC 59
           ||| | ||||||||||||||||||||||||||||||||||||||||||||||||| |||
Qy       4 GACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACAC 63

Db      60 AATGACTTGCAGGGCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGG 119
           |||||||||||||||| | |||||||||||||||||||||||| ||||||||||||||
Qy      64 AATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGG 123

Db     120 ATCCTCCCCCAAACCCTGGATTTATGCCACATCCAACCTGGCTTCTGGAGTCCCTGCTCG 179
           |||||||||||||| |||||||| ||||||||||||:||||||||||||||||||||||
Qy     124 ATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCG 183

Db     180 CTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCATCAGAGTGGAGGCTGA 239
           |||||||||| |||||||||||||||||||||||||||| ||| ||||||||||||||
Qy     184 CTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGA 243

Db     240 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTAGTAACCCGCTCACGTTCGGTGCTGG 299
           |||||||||||||||||||||||||||||||| ||| ||| | ||||||||| |  ||
```

```
Qy    244 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGG 303

Db    300 GACCAAGCTGGAGAT 314
          | ||||||||||| ||
Qy    304 GTCCAAGCTGGAAAT 318


RESULT   13
LOCUS       S48339        312 bp    mRNA            ROD       14-JAN-1993
DEFINITION  Ig V kappa -anti-idiotypic Fab [mice, mRNA Partial, 312 nt].
ACCESSION   S48339
NID         g257749
KEYWORDS    .
SOURCE      Mus sp.
  ORGANISM  Mus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 312)
  AUTHORS   Kasai,Y., Herlyn,D., Sperlagh,M., Maruyama,H., Matsushita,S. and
            Linnenbach,A.J.
  TITLE     Molecular cloning of murine monoclonal anti-idiotypic Fab
  JOURNAL   J. Immunol. Methods 155 (1), 77-89 (1992)
  MEDLINE   93017981
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 118382] from the original journal article.
            This sequence comes from Fig. 5.
FEATURES             Location/Qualifiers
     source          1. .312
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
     gene            1. .312
                     /partial
                     /note="anti-idiotypic Fab"
                     /gene="Ig V<down>&kgr;</down>"
     CDS             1. .312
                     /partial
                     /gene="Ig V<kappa>"
                     /note="anti-idiotypic Fab; This sequence comes from Fig.
                     5"
                     /codon_start=1
                     /db_xref="PID:g257750"
                     /translation="VMTQSPAILSASPGEKVTMTCRASSSVSYMHWYQQKPGSSPKPW
                     IYATSNLASGVPARFSGSGSGTSYSLTISRVEAEDAATYYCQQWSSNPWTFGGGTKLE
                     IK"
BASE COUNT       76 a     88 c     80 g     68 t
ORIGIN

  Query Match           37.6%;  Score 268;  DB 28;  Length 312;
  Best Local Similarity  93.9%;  Pred. No. 7.55e-198;
  Matches   290;  Conservative    1;  Mismatches  17;  Indels   1;  Gaps   1;

Db      5 TGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CACAATGACTTGC 63
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Qy     14 TGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGC 73

Db     64 AGGGCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGGATCCTCCCCC 123
          ||||||| | |||||||||||||||||||||||| ||||||||||||||||||||||||
Qy     74 AGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCC 133

Db    124 AAACCCTGGATTTATGCCACATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGC 183
          ||||| |||||||||| ||||||||||||:|||||||||||||||||||||||||||||||
Qy    134 AAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGC 193

Db    184 AGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCC 243
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Qy    194 GGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCC 253

Db    244 ACTTATTACTGCCAGCAGTGGAGTAGTAACCCATGGACGTTCGGTGGAGGCACCAAGCTG 303
          |||||||||||||||||||||||| ||| ||||   |||||||| || ||  ||||||||
Qy    254 ACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTG 313

Db    304 GAAATCAAA 312
          ||||| |||
Qy    314 GAAATAAAA 322


RESULT   14
LOCUS       AF056217      318 bp    mRNA            ROD       13-APR-1998
DEFINITION  Mus musculus monoclonal antibody aH7:38 IgG1 light chain mRNA,
            partial cds.
ACCESSION   AF056217
NID         g3044131
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 318)
  AUTHORS   Sandstrom,P., Johansson,A., Ullen,A., Behravan,G. and Stigbrand,T.
  TITLE     Differences in sequence and affinity between three monoclonal
            anti-idiotypic antibodies against one anti-placental alkaline
            phosphatase idiotypic antibody
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 318)
  AUTHORS   Sandstrom,P., Johansson,A., Ullen,A., Behravan,G. and Stigbrand,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-1998) Immunology, Umea University, Umea 90185,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .318
                     /organism="Mus musculus"
                     /strain="Balb/c"
                     /db_xref="taxon:10090"
                     /cell_line="hybridoma"
     CDS             <1. .>318
                     /note="anti-idiotypic antibody against anti-placental
                     alkaline phosphatase antibody; variable region"
                     /codon_start=1
                     /product="monoclonal antibody aH7:38 IgG1 light chain"
                     /db_xref="PID:g3044132"
                     /translation="DIQLTQSPAILSASPGEKVSMTCRASSSVSYIHWYQQKPGSSPK
                     PWIYATSNLASGVPVRFSGSGSGTSYSLTISRVEAEDAATYYCQQWSGNPLTFGAGTK
                     LEIK"
BASE COUNT       74 a     91 c     81 g     72 t
ORIGIN

  Query Match           37.6%;  Score 268;  DB 28;  Length 318;
  Best Local Similarity  92.5%;  Pred. No. 7.55e-198;
  Matches   295;  Conservative    1;  Mismatches  22;  Indels   1;  Gaps   1;

Db      1 GACATTCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CTC 59
          ||| | |||||||||||||||||||||||||||||||||||||||||||||||||| | |
Qy      4 GACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACAC 63

Db     60 AATGACTTGCAGGGCCAGCTCAAGTGTAAGTTACATACACTGGTACCAGCAGAAGCCAGG 119
          |||||||||||||||| | |||||||||||||||| |||||||| ||||||||||||||
Qy     64 AATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGG 123

Db    120 ATCCTCCCCCAAACCCTGGATTTATGCCACATCCAACCTGGCTTCTGGAGTCCCTGTTCG 179
          |||||||||||||| |||||||| ||||||||||||:|||||||||||||||| |||
Qy    124 ATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCG 183

Db    180 CTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGA 239
          |||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||
Qy    184 CTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGA 243

Db    240 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTGGTAACCCGCTCACGTTCGGTGCTGG 299
          |||||||||||||||||||||||||||||||||||| ||| ||| | |||||||||| |  ||
Qy    244 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGG 303

Db    300 GACCAAGCTGGAGATCAAA 318
          | ||||||||| || |||
Qy    304 GTCCAAGCTGGAAATAAAA 322
```

```
RESULT   15
LOCUS       MUSIGKAAR    1370 bp    DNA             ROD       05-FEB-1992
DEFINITION  Mouse Ig kappa active V-region from 70Z/3 cells.
ACCESSION   K01641
NID         g196460
KEYWORDS    C-region; V-region; immunoglobulin light chain;
            immunoglobulin-kappa.
SOURCE      Mouse 70Z/3 cell DNA.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 1370)
  AUTHORS   Parslow,T.G., Blair,D.L., Murphy,W.J. and Granner,D.K.
  TITLE     Structure of the 5' ends of immunoglobulin genes: A novel conserved
            sequence
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 81, 2650-2654 (1984)
  MEDLINE   84194064
COMMENT     Corrections to the sequence were sent by Dr. Richard E. Manroe and
            Dr. Tristram G. Parslow on Dec. 9, 1991.
FEATURES             Location/Qualifiers
     source          1. .1370
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     exon            <776. .824
                     /note="Ig kappa chain variable region"
                     /number=1
     CDS             join(776. .824,1003. .1337)
                     /partial
                     /note="Ig kappa chain variable region"
                     /codon_start=1
                     /db_xref="PID:g196461"
                     /translation="MDFQVQIFSFLLISASVIMSRGQIVLSQSPAILSASPGEKVTMT
                     CRASSSVSYMHWYQQKPGSSPKPWIYATSNLASGVPARFSGSGSGTSYSLTISRVEAE
                     DAATYYCQQWSSNPRTFGGGTKLEIK"
     intron          825. .1002
                     /note="kappa cds intron"
     exon            1003. .>1337
                     /note="Ig kappa chain variable region"
                     /number=2
     intron          1338. .>1369
                     /note="kappa cds intron J-C"
BASE COUNT      404 a    273 c    253 g    440 t
ORIGIN

  Query Match           37.6%;  Score 268;  DB 28;  Length 1370;
  Best Local Similarity  93.9%;  Pred. No. 7.55e-198;
  Matches   290;  Conservative    1;  Mismatches  17;  Indels   1;  Gaps   1;

Db   1033 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CACAATGACTTGCAGG 1091
            |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db   1092 GCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGGATCCTCCCCCAAA 1151
            |||| | |||||||||||||||||||||||| |||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db   1152 CCATGGATTTATGCCACATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGT 1211
            || ||||||||| ||||||||||||:||||||||||||||||||||||||||||||| ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db   1212 GGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 1271
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db   1272 TATTACTGCCAGCAGTGGAGTAGTAACCCACGGACGTTCGGTGGAGGCACCAAGCTGGAA 1331
            |||||||||||||||||||| ||| |||||  |||||||| || ||  ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAA 316

Db   1332 ATCAAACGT 1340
            || ||| ||
Qy    317 ATAAAAGGT 325
```

Search completed: Sat Jan 9 09:50:43 1999
Job time : 1060 secs.

****

****

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Jan 9 12:36:16 1999; MasPar time 116.06 Seconds
834.616 Million cell updates/sec

Tabular output not generated.

Title: >US-08-704-178-1
Description: (1-712) from US08704178.seq (1 of 4)
Perfect Score: 712
N.A. Sequence: 1 ATGGACCTGCAGCTGACCCA.....GGACCACGGTCACCGTCTCC 712
Comp: TACCTGGACGTCGACTGGGT.....CCTGGTGCCAGTGGCAGAGG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 8.732; Variance 5.071; scale 1.722

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	705	99.0	711	29	T65006	Single-chain anti-erb	0.00e+00
2	701	98.5	711	9	Q55180	Sequence encoding the	0.00e+00
3	699	98.2	711	18	T17728	Anti-erbB2 scFv cDNA.	0.00e+00
4	410	57.6	720	34	T91615	cDNA encoding an anti	1.02e-272
5	375	52.7	720	29	T65007	Single-chain anti-erb	4.34e-247
6	375	52.7	720	9	Q55181	Sequence encoding the	4.34e-247
7	363	51.0	732	12	Q73678	Fv(GP-2) immunosuppre	2.58e-238
8	279	39.2	729	12	Q73679	Fv(GP-4) immunosuppre	4.18e-177
9	266	37.4	462	21	T31541	3H1 heavy chain varia	1.13e-167
10	266	37.4	462	38	T99435	Anti-idiotype antibod	1.13e-167
11	266	37.4	462	22	T34542	Monoclonal anti-idiot	1.13e-167
12	265	37.2	384	11	Q65630	Murine variable regio	5.98e-167
13	264	37.1	9208	11	Q65629	Vector contg. TCAE 8	3.17e-166
14	263	36.9	318	34	T79900	Anti-Factor IX MAb ch	1.68e-165
15	263	36.9	321	34	T77377	Mouse anti-human Fact	1.68e-165
16	263	36.9	335	34	T79899	Anti-Factor IX MAb BC	1.68e-165
17	263	36.9	403	1	N91147	2H7 Vh sequence.	1.68e-165
18	263	36.9	403	2	N70972	2H7 VL sequence in wh	1.68e-165
19	263	36.9	426	31	T70869	2H7 light chain varia	1.68e-165
20	263	36.9	426	30	T51043	Coding sequence for l	1.68e-165
21	263	36.9	426	32	T36317	2H7 antibody light ch	1.68e-165
22	263	36.9	426	40	V18594	Mouse 2H7 antibody li	1.68e-165
23	263	36.9	426	40	V18558	Mouse 2H7 antibody li	1.68e-165
24	263	36.9	426	40	V03927	Mouse 2H7 antibody li	1.68e-165
25	262	36.8	321	6	Q39412	Mab32 light chain V-g	8.93e-165
26	258	36.2	387	4	Q27350	Encodes A5B7 antibody	7.05e-162
27	258	36.2	417	7	Q43843	Chimeric 128.1 VH, mo	7.05e-162
28	258	36.2	10704	7	Q43846	Plasmid pAH4625.	7.05e-162
29	258	36.2	10844	7	Q43848	Plasmid pAH4808.	7.05e-162
30	258	36.2	11529	7	Q43844	Plasmid pAH4602.	7.05e-162
31	258	36.2	12132	7	Q43847	Plasmid pAH4807.	7.05e-162
32	257	36.1	387	3	Q20983	Encodes Variable regi	3.74e-161
33	256	36.0	797	22	T29058	CC49/212 SCA PLAP DNA	1.98e-160
34	256	36.0	803	22	T29059	PLAP CC49/212 SCA DNA	1.98e-160
35	256	36.0	1460	7	Q43291	Sequence encoding div	1.98e-160
36	255	35.8	731	7	Q43287	Sequence encoding a m	1.05e-159
37	255	35.8	732	24	T42508	Murine A5B57 Light ch	1.05e-159
38	253	35.5	435	6	Q36613	Anti-IL2R-alpha antib	2.94e-158
39	253	35.5	435	3	Q15115	IL-2 chimeric antibod	2.94e-158
40	251	35.3	292	29	T63502	Monoclonal antibody A	8.23e-157
41	251	35.3	292	24	T36662	Monoclonal antibody A	8.23e-157
42	251	35.3	292	1	Q06213	VK domain of antibody	8.23e-157
43	249	35.0	443	11	Q45428	KM-750 heavy chain.	2.30e-155
44	248	34.8	321	1	N91661	Chimeric antibody lig	1.22e-154
45	247	34.7	420	2	Q11970	Sequence encoding lig	6.44e-154

ALIGNMENTS

RESULT 1
ID T65006 standard; cDNA; 711 BP.
AC T65006;
DT 05-JUN-1997 (first entry)
DE Single-chain anti-erbB2 antibody e23(Fv) cDNA.
KW Single chain antibody; variable region; light chain; heavy chain;
KW breast cancer; ovarian cancer; non-small cell lung carcinoma;
KW immunodiagnosis; treatment; cytotoxic agent; erbB-2; ds.
OS Mus musculus.
OS Synthetic.
FH Key Location/Qualifiers
FT mat_peptide 1..711
FT /*tag= a
FT /product= e23(Fv)
PN US5587458-A.
PD 24-DEC-1996.
PF 07-OCT-1991; 772270.
PR 07-OCT-1991; US-772270.
PR 30-JUN-1992; US-906555.
PR 14-MAY-1993; US-061092.
PA (ARON-) ARONEX PHARM INC.
PI Bird RE, Kasprzyk PG, King CR;
DR WPI; 97-064831/06.
DR P-PSDB; W15185.
PT Single chain antibodies specific for erbB-2 protein, gp185 - with
PT labels or cytotoxin, useful for detection and treatment of tumour
PT cells expressing this protein
PS Example 8; Columns 25-28; 28pp; English.
CC The present cDNA sequence codes for a claimed single-chain antibody,
CC designated e23(Fv), which binds to erbB-2. Monoclonal antibody e23
CC was generated by immunising mice with N/erbB-2 cells overexpressing
CC the gp185 protein, removing spleen cells and producing hybridomas
CC by standard techniques. Messenger RNA coding for the anti-erbB-2
CC monoclonal antibody was isolated and converted to cDNA. Regions
CC coding for the heavy- and light- chain variable regions were then
CC amplified by PCR and joined via a sequence encoding a peptide
CC linker. The resulting single-chain antibody is useful for in vitro

```
CC   diagnosis of tumour cells which overexpress the erbB-2 gp185
CC   marker, e.g. breast, ovarian and non-small cell lung carcinomas,
CC   and, when coupled to a cytotoxic agent, to treat such tumours.
SQ   Sequence   711 BP;    175 A;    182 C;    189 G;    164 T;

  Query Match           99.0%;  Score 705;  DB 29;  Length 711;
  Best Local Similarity  99.9%;  Pred. No. 0.00e+00;
  Matches   711;  Conservative    0;  Mismatches   0;  Indels   1;  Gaps   1;

Db      1 atggacctgcagctgacccagtctccagcaatcctgtctgcatctccaggggagaaggt- 59
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db     60 cacaatgacttgcagggccaccccaagtgtaagttacatgcactggtatcagcagaagcc 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 120

Db    120 aggatcctcccccaaaccttggatttataccacatccaacctkgcttctggagtccctgc 179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180

Db    180 tcgcttcagtggcggtgggtctgggacctcttactctctcacagtcagcagagtggaggc 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db    240 tgaagatgctgccacttattactgccagcagtggagtcgtagcccacccacgttcggagg 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 300

Db    300 ggggtccaagctggaaataaaaggttctacctctggttctggtaaatcttctgaaggtaa 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301 GGGGTCCAAGCTGGAAATAAAAGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAA 360

Db    360 aggtgtgcagctgcaggagtcaggacctgaggtggtgaagcctggaggttcaatgaagat 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361 AGGTGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGAT 420

Db    420 atcctgcaagacttctggttactcattcactggccacaccatgaactgggtgaagcagag 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421 ATCCTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAG 480

Db    480 ccatggaaagaaccttgagtggattggacttattaatccttacaatggtgatactaacta 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    481 CCATGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTA 540

Db    540 caaccagaagttcaagggcaaggccacatttactgtagacaagtcgtccagcacagccta 599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541 CAACCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTA 600

Db    600 catggagctcctcagtctgacatctgaggactctgcagtctattactgtgcaaggagggt 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    601 CATGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGT 660

Db    660 tacggactggtacttcgatgtctggggcgcagggaccacggtcaccgtctcc 711
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    661 TACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTCC 712


RESULT    2
ID   Q55180 standard; cDNA; 711 BP.
AC   Q55180;
DT   21-JUL-1994  (first entry)
DE   Sequence encoding the single chain anti-erbB2 antibody, Ab no.23.
KW   Single chain anti-erbB1 antibody; cancer therapy; prevention;
KW   monoclonal antibody; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   cds             1..171
FT                   /*tag=  a
PN   WO9400136-A.
PD   06-JAN-1994.
PF   21-OCT-1992; U08545.
PR   30-JUN-1992; US-906555.
PA   (MOLE-) MOLECULAR ONCOLOGY INC.
PI   Kasprzyk PG,  King CR;
DR   WPI; 94-025878/03.
DR   P-PSDB; R45442.
PT   Treatment of malignancies over-expressing ERB-[2 - using at least
PT   2 monoclonal antibodies which recognise different epitopes on
PT   gp185
PS   Example; Fig 7; 37pp; English.
CC   The source of human erbB-2 protein for the prodn. of antibodies no.
CC   23 and 21 (Q55180/R45442; Q55181/R45443) is a NIH/3T3 cell
CC   engineered to express the human erbB-2 protein on its surface (N/
CC   erbB-2). Abs no. 21 and 23 are directed against the extracellular
CC   domain of gp185 erbB-2. Nude mice manipulated to produce rapidly
CC   growing tumours were used in a trial of the efficacy of the Abs. In
CC   animals given a combination of the 2 Abs, tumours completely
CC   regressed after 11 days.
SQ   Sequence   711 BP;    175 A;    183 C;    187 G;    165 T;

  Query Match           98.5%;  Score 701;  DB 9;  Length 711;
  Best Local Similarity  99.6%;  Pred. No. 0.00e+00;
  Matches   709;  Conservative    0;  Mismatches   2;  Indels   1;  Gaps   1;

Db      1 atggacctgcagctgacccagtctccagcaatcctgtctgcatctccaggggagaaggt- 59
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db     60 cacaatgacttgcagggccaccccaagtgtaagttacatgcactggtatcagcagaagcc 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 120

Db    120 aggatcctcccccaaaccttggatttataccacatccaacctkgcttctggagtccctgc 179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180

Db    180 tcgcttcagtggcggtgggtctgggacctcttactctctcacagtcagcagagtggaggc 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db    240 tgaagatgctgccacttattactgccagcagtggagtcgtagcccacccacgttcggagg 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 300

Db    300 ggggtccaagctggaaataaaaggttctacctctggttctggtaaatcttctgaaggtaa 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301 GGGGTCCAAGCTGGAAATAAAAGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAA 360

Db    360 aggtgtgcagctgcaggattcaggacctgaggtggtgaagcctggaggttcaatgaagat 419
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Qy    361 AGGTGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGAT 420

Db    420 atcctgcaagacttctggttactcattcactggccacaccatgaactgggtgaagcagag 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421 ATCCTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAG 480

Db    480 ccatggaaagaaccttgagtggattggacttattaatccttacaatggtgatactaacta 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    481 CCATGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTA 540

Db    540 caaccagaagttcaagggcaaggccacatttactgtagacaagtcgtccagcacagccta 599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541 CAACCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTA 600

Db    600 catggagctcctcagtctgacatctgaggactctgcactctattactgtgcaaggagggt 659
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Qy    601 CATGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGT 660

Db    660 tacggactggtacttcgatgtctggggcgcagggaccacggtcaccgtctcc 711
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    661 TACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTCC 712


RESULT    3
ID   T17728 standard; cDNA; 711 BP.
```

```
AC   T17728;
DT   21-MAY-1996  (first entry)
DE   Anti-erbB2 scFv cDNA.
KW   Oncoprotein; erbB2; cell proliferation; tumour; cancer;
KW   intracellular antibody homologue; single chain antibody; scFv;
KW   gene therapy; ds.
OS   Synthetic.
PN   WO9607321-A1.
PD   14-MAR-1996.
PF   23-AUG-1995; U10740.
PR   06-SEP-1994; US-301339.
PR   06-JUN-1995; US-468252.
PA   (UABR-) UAB RES FOUND.
PI   Curiel DT,  Deshane J;
DR   WPI; 96-171307/17.
DR   P-PSDB; R94020.
PT   Inhibition of proliferation or survival of, esp. malignant erbB2,
PT   cells - by introducing nucleic acid mol. encoding antibody homologue
PT   which is expressed and binds, pref. erbB2, protein intracellularly
PS   Claim 42; Page 29-30; 48pp; English.
CC   A nucleic acid comprises a first sequence encoding a signal peptide
CC   (R94019) linked to a second sequence (T17728) encoding a single
CC   chain Fv fragment (R94020) that binds a human erbB2 oncoprotein.
CC   The anti-erbB2 sFv portion is obtained by PCR using e23scFv
CC   plasmid as template.  The signal peptide directs the scFv to the
CC   endoplasmic reticulum.  The nucleic acid is incorporated into a
CC   plasmid or viral vector to facilitate expression of the scFv antibody
CC   homologue within e.g. an epithelial carcinoma cell.  Intracellular
CC   expression of the homologue inhibits surface expression of erbB2 and
CC   thereby inhibits cell proliferation and cell survival and decreases
CC   tumorigenicity.
SQ   Sequence   711 BP;    175 A;    183 C;    189 G;    164 T;

  Query Match           98.2%;  Score 699;  DB 18;  Length 711;
  Best Local Similarity  99.3%;  Pred. No. 0.00e+00;
  Matches   707;  Conservative    1;  Mismatches   3;  Indels   1;  Gaps   1;

Db      1 atggacctgcagctgacccagtctccagcaatcctctctgcatctccaggggagaaggt- 59
          ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db     60 cacaatgacttgcagggccaccccaagtgtaagttacatgcactggtatcagcagaagcc 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 120

Db    120 aggatcctcccccaaaccttggatttataccacatccaacctggcttctggagtccctgc 179
          |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Qy    121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180

Db    180 tcgcttcagtggcggtgggtctgggacctcttactctctcacagtcagcagagtggaggc 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db    240 tgaagatgctgccacttattactgccagcagtggagtcgtagcccacccacgttcggagg 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 300

Db    300 ggggtccaagctggaaataaaaggttctacctctggttctggtaaatcttctgaaggtaa 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301 GGGGTCCAAGCTGGAAATAAAAGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAA 360

Db    360 aggtgtgcagctgcaggagtcaggacctgaggtggtcaagcctggaggttcaatgaagat 419
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Qy    361 AGGTGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGAT 420

Db    420 atcctgcaagacttctggttactcattcactggccacaccatgaactgggtgaagcagag 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421 ATCCTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAG 480

Db    480 ccatggaaagaaccttgagtggattggacttattaatccttacaatggtgatactaacta 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    481 CCATGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTA 540

Db    540 caaccagaagttcaagggcaaggccacatttactgtagacaagtcgtccagcacagccta 599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541 CAACCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTA 600

Db    600 catggagctcctcagtctgacatctgaggactctgcagtctattactgtgcaaggagggt 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    601 CATGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGT 660

Db    660 tacggactggtacttcgatgtctgggggcagggaccacggtcaccgtctcc 711
          ||||||||||||||||||||||||||| ||||||||||||||||||||||
Qy    661 TACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTCC 712


RESULT    4
ID   T91615 standard; cDNA to mRNA; 720 BP.
AC   T91615;
DT   22-DEC-1997  (first entry)
DE   cDNA encoding an anti-gp130 antibody derived scFv.
KW   Single chain Fv; scFv; antibody; E. coli; anti-T3 antibody;
KW   anti-gp130 antibody; inclusion body; chaperonin; ds.
OS   Synthetic.
PN   J09220092-A.
PD   26-AUG-1997.
PF   15-FEB-1996; 027622.
PR   15-FEB-1996; JP-027622.
PA   (TOYJ ) TOSOH CORP.
DR   WPI; 97-474306/44.
DR   P-PSDB; W25784.
PT   Producing single chain Fv antibody in Escherichia coli - by
PT   expression in an inclusion body, followed by protein folding or by
PT   co-expression with a chaperonin as a soluble fraction
PS   Claim 4; Page 6-7; 9pp; Japanese.
CC   The sequences given in T91614-15 encode single chain Fv (scFv)
CC   antibodies which are produced in E. coli.  The scFV's are derived
CC   from an anti-T3 antibody or an anti-gp130 antibody and are produced
CC   either by: (1) expression as an inclusion body, followed by folding
CC   (i.e. by denaturation and solubilisation) or (2) expression as a
CC   soluble fraction by co-expression with chaperonin.  Using the methods,
CC   scFv can be produced in E. coli, either in a soluble fraction or in
CC   inclusion bodies.  Like the parent double chain antibodies, the scFv
CC   polypeptides are useful in immunological diagnosis or for separation
CC   and purification, but they lessen the side effects caused by antibody
CC   constant regions.  The availability of scFv polypeptides contributes
CC   to fundamental research and development of diagnostic and therapeutic
CC   drugs.
SQ   Sequence   720 BP;    169 A;    193 C;    190 G;    168 T;

  Query Match           57.6%;  Score 410;  DB 34;  Length 720;
  Best Local Similarity  83.1%;  Pred. No. 1.02e-272;
  Matches   537;  Conservative    1;  Mismatches 104;  Indels   4;  Gaps   3;

Db      9 gctcacccagtctccagcattgatatctgcatctccaggggagaaggt-caccatgacct 67
          ||| ||||||||||||||| |  | |||||||||||||||||||||| ||| ||||| |
Qy     12 GCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTT 71

Db     68 gcaatgtcagctcaagtgttacttccatgtattggtaccagcagaagccaggatcctccc 127
          |||  | || | ||||||| | || |||| | ||||| ||||||||||||||||||||
Qy     72 GCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCC 131

Db    128 ccaaaccctggatttatctcacatccaacctggcttctggagtccctcctcgctccagtg 187
          ||||||| |||||||||  ||||||||||||:|||||||||||||||| |||||| |||||
Qy    132 CCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTG 191

Db    188 gcagtgggtctgggaccacttactctctcacaatcagcagcttggaggctgaagatgctg 247
          || |||||||||||||| |||||||||||||| |||||||  ||||||||||||||||||
Qy    192 GCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTG 251

Db    248 ccacttattactgccagcagtggagtactaacccgctcacgttcggtgctgggaccaagc 307
          |||||||||||||||||||||||||||  || ||| | |||||||||| |  ||| ||||||
Qy    252 CCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGC 311

Db    308 tcgagctgggtggcggtggctcgggcggtggtgggtcgggtggcggcggatcggacgtcc 367
          | ||  |    ||   | ||| ||   ||||   ||   ||   ||   |  ||  || |
```

```
Qy    312 TGGAAATAAAAGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAA--GGT-GTGC 368

Db    368 aattgcagcagtctggacctgaactggtgaagcctggggcttcagtgaagataccctgca 427

Qy    369 AGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCCTGCA 428

Db    428 aggcttcaggatacacattcactgactacaacatggactgggtgaagcagagccatggaa 487

Qy    429 AGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCATGGAA 488

Db    488 agagccttgagtggattggagatattaattctcatagtggtggtattatctacaaccaaa 547

Qy    489 AGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAACCAGA 548

Db    548 agttcaaggacaaggccatattgactgtagataagtcctccagcacagcctacatggagc 607

Qy    549 AGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATGGAGC 608

Db    608 tccgcagcctgacatctgaggacactgcagtctattattgtgcaag 653

Qy    609 TCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654


RESULT    5
ID   T65007 standard; cDNA; 720 BP.
AC   T65007;
DT   05-JUN-1997  (first entry)
DE   Single-chain anti-erbB2 antibody e21(Fv) cDNA.
KW   Single chain antibody; variable region; light chain; heavy chain;
KW   breast cancer; ovarian cancer; non-small cell lung carcinoma;
KW   immunodiagnosis; treatment; cytotoxic agent; erbB-2; ds.
OS   Mus musculus.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     1..720
FT                   /*tag= a
FT                   /product= e21(Fv)
PN   US5587458-A.
PD   24-DEC-1996.
PF   07-OCT-1991; 772270.
PR   07-OCT-1991; US-772270.
PR   30-JUN-1992; US-906555.
PR   14-MAY-1993; US-061092.
PA   (ARON-) ARONEX PHARM INC.
PI   Bird RE,  Kasprzyk PG,  King CR;
DR   WPI; 97-064831/06.
DR   P-PSDB; W15186.
PT   Single chain antibodies specific for erbB-2 protein, gp185  - with
PT   labels or cytotoxin, useful for detection and treatment of tumour
PT   cells expressing this protein
PS   Example 9; Columns 27-30; 28pp; English.
CC   The present cDNA sequence codes for a claimed single-chain antibody,
CC   designated e21(Fv), which binds to erbB-2. Monoclonal antibody e21
CC   was generated by immunising mice with N/erbB-2 cells overexpressing
CC   the gp185 protein, removing spleen cells and producing hybridomas
CC   by standard techniques. Messenger RNA coding for the anti-erbB-2
CC   monoclonal antibody was isolated and converted to cDNA. Regions
CC   coding for the heavy- and light- chain variable regions were then
CC   amplified by PCR and joined via a sequence encoding a peptide
CC   linker. The resulting single-chain antibody is useful for in vitro
CC   diagnosis of tumour cells which overexpress the erbB-2 gp185
CC   marker, e.g. breast, ovarian and non-small cell lung carcinomas,
CC   and, when coupled to a cytotoxic agent, to treat such tumours.
SQ   Sequence   720 BP;    184 A;    178 C;    179 G;    179 T;

  Query Match           52.7%;  Score 375;  DB 29;  Length 720;
  Best Local Similarity  79.3%;  Pred. No. 4.34e-247;
  Matches   511;  Conservative    1;  Mismatches 131;  Indels   1;  Gaps   1;

Db      2 tgcagctgacccagtctccagcaatcatgtctgcatctccaggggaaaaggt-caccatg 60

Qy      8 TGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATG 67

Db     61 acctgcagtgccagctcaagtgtaagtaacatgcactggtatcagcagaagtcaagcacc 120

Qy     68 ACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCC 127

Db    121 tcccccaaactctgggtttatgacacatccaaactggcttctggagtcccaggtcgcttc 180

Qy    128 TCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTC 187

Db    181 agtggcagtgggtctggaaactcttactctctcacgatcagcagcatggaggctgaagat 240

Qy    188 AGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGAT 247

Db    241 gctgccacttattattgttatcaggggagtgggtacccattcacgttcggctcggggaca 300

Qy    248 GCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCC 307

Db    301 aagttggaaataaaaggttctacctccggatctggtaaatcttctgaaggtaaaggtgtg 360

Qy    308 AAGCTGGAAATAAAAGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGGTGTG 367

Db    361 cagctgcagcagtctgggggttgagcttgtccgaggaggggccttagtcaagttgtcctgc 420

Qy    368 CAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCCTGC 427

Db    421 aaagcttctgacttcaacattaaagactattatatccactgggtgaagcagaggcctgaa 480

Qy    428 AAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCATGGA 487

Db    481 cagggcctggaatggattggatggattcatcctgagaatggtaatactgtatatgacccg 540

Qy    488 AAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAACCAG 547

Db    541 aaattccagggcaaggccagtataacagcagacacatcctccaacgcggcctaccttcag 600

Qy    548 AAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATGGAG 607

Db    601 ctcagcagcctgacatctgaggacactgccgtctattactgtgc 644

Qy    608 CTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGC 651


RESULT    6
ID   Q55181 standard; cDNA; 720 BP.
AC   Q55181;
DT   21-JUL-1994  (first entry)
DE   Sequence encoding the single chain anti-erbB2 antibody, Ab no.21.
KW   Single chain anti-erbB1 antibody; cancer therapy; prevention;
KW   monoclonal antibody; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   cds             1..171
FT                   /*tag= a
PN   WO9400136-A.
PD   06-JAN-1994.
PF   21-OCT-1992; U08545.
PR   30-JUN-1992; US-906555.
PA   (MOLE-) MOLECULAR ONCOLOGY INC.
PI   Kasprzyk PG,  King CR;
DR   WPI; 94-025878/03.
DR   P-PSDB; R45443.
PT   Treatment of malignancies over-expressing ERB-[2 - using at least
PT   2 monoclonal antibodies which recognise different epitopes on
PT   gp185
PS   Example; Fig 8; 37pp; English.
CC   The source of human erbB-2 protein for the prodn. of antibodies no.
CC   23 and 21 (Q55180/R45442; Q55181/R45443) is a NIH/3T3 cell
CC   engineered to express the human erbB-2 protein on its surface (N/
CC   erbB-2). Abs no. 21 and 23 are directed against the extracellular
CC   domain of gp185 erbB-2. Nude mice manipulated to produce rapidly
CC   growing tumours were used in a trial of the efficacy of the Abs. In
CC   animals given a combination of the 2 Abs, tumours completely
CC   regressed after 11 days.
SQ   Sequence   720 BP;    184 A;    178 C;    179 G;    179 T;
```

```
  Query Match           52.7%;  Score 375;  DB 9;  Length 720;
  Best Local Similarity  79.3%;  Pred. No. 4.34e-247;
  Matches   511;  Conservative    1;  Mismatches 131;  Indels   1;  Gaps   1;

Db      2 tgcagctgacccagtctccagcaatcatgtctgcatctccaggggaaaaggt-caccatg 60
          |||||||||||||||||||||||||| |||||||||||||||||| ||||| ||| |||
Qy      8 TGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATG 67

Db     61 acctgcagtgccagctcaagtgtaagtaacatgcactggtatcagcagaagtcaagcacc 120
          || ||||| |||| | |||||||||| |||||||||||||||||||| || | ||
Qy     68 ACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCC 127

Db    121 tcccccaaactctgggtttatgacacatccaaactggcttctggagtcccaggtcgcttc 180
          |||||||||| ||| ||||| ||||||||| ||:||||||||||||||| | |||||||
Qy    128 TCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTC 187

Db    181 agtggcagtgggtctggaaactcttactctctcacgatcagcagcatggaggctgaagat 240
          |||||| |||||||||| | |||||||||||||||  |||||||  ||||||||||||||
Qy    188 AGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGAT 247

Db    241 gctgccacttattattgttatcaggggagtgggtacccattcacgttcggctcggggaca 300
          |||||||||||||| ||  | ||| ||||| |   ||||  |||||||||   |||| |
Qy    248 GCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCC 307

Db    301 aagttggaaataaaaggttctacctccggatctggtaaatcttctgaaggtaaaggtgtg 360
          ||| |||||||||||||||||||||| || ||||||||||||||||||||||||||||||
Qy    308 AAGCTGGAAATAAAAGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGGTGTG 367

Db    361 cagctgcagcagtctggggttgagcttgtccgaggaggggccttagtcaagttgtcctgc 420
          |||||||| |||| ||   |||| | ||       || |  | | | ||| | ||||||
Qy    368 CAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCCTGC 427

Db    421 aaagcttctgacttcaacattaaagactattatatccactgggtgaagcagaggcctgaa 480
          ||  ||||||  | |    | |  | | |    ||  ||||||||||||||| | || |
Qy    428 AAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCATGGA 487

Db    481 cagggcctggaatggattggatggattcatcctgagaatggtaatactgtatatgacccg 540
           ||  ||| || |||||||||   ||| ||||| | |||||| |||||   ||  ||| |
Qy    488 AAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAACCAG 547

Db    541 aaattccagggcaaggccagtataacagcagacacatcctccaacgcggcctaccttcag 600
          || ||| ||||||||||||   | || | |||||  || |||| | | |||||| |  ||
Qy    548 AAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATGGAG 607

Db    601 ctcagcagcctgacatctgaggacactgccgtctattactgtgc 644
          |||  ||| ||||||||||||||| |||| ||||||||||||||
Qy    608 CTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGC 651


RESULT    7
ID   Q73678 standard; DNA; 732 BP.
AC   Q73678;
DT   09-MAY-1995  (first entry)
DE   Fv(GP-2) immunosuppressive.
KW   MAb; monoclonal antibody; hybridoma; interleukin-2; IL-2;
KW   Fv; antibody variable region; GP-2; Fv(GP-2);
KW   immunosuppressive; ss.
OS   Mus sp.
PN   EP-621338-A.
PD   26-OCT-1994.
PF   21-APR-1994; 106257.
PR   21-APR-1993; JP-094491.
PR   07-MAR-1994; JP-036065.
PA   (AJIN ) AJINOMOTO KK.
PI   Hamura J,  Kanayama Y,  Nakazawa H,  Shimamura T,  Sugamura K;
PI   Takeshita T;
DR   WPI; 94-325948/41.
DR   P-PSDB; R60780.
PT   Immunosuppressant polypeptide - has ability to block
PT   interleukin-2 response
PS   Claim 17; Page 29; 37pp; English.
CC   MAb capable of binding to the gamma chain of the IL-2 receptor, and
CC   thus of blocking the IL-2 response, is produced by mouse hybridoma
CC   line GP-2 (FERM BP-4641).  DNA encoding the variable region of
CC   this MAb was expressed in E. coli, yielding Fv(GP-2) with
CC   immumosuppressive activity.
SQ   Sequence   732 BP;    180 A;    204 C;    182 G;    166 T;

  Query Match           51.0%;  Score 363;  DB 12;  Length 732;
  Best Local Similarity  79.8%;  Pred. No. 2.58e-238;
  Matches   550;  Conservative    1;  Mismatches 128;  Indels  10;  Gaps   4;

Db      1 atggacatcctgctgacccagtctccatcaatcatgtctgcatctctaggggaacgggt- 59
          |||||| | | |||||||||||||||| ||||| |||||||||||| ||||||   |||
Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db     60 caccatgacctgcactgccagctcaagtgtaagttccagttacttgcactggtaccagca 119
          ||| ||||| ||||  |||| | |||||||||||| ||      |||||||||| |||||
Qy     61 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACA------TGCACTGGTATCAGCA 114

Db    120 gaagccaggatcctcccccaaactctggatttatagcacatccaacctggcttctggagt 179
          |||||||||||||||||||||  |||||||||| ||||||||||:||||||||||||
Qy    115 GAAGCCAGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGT 174

Db    180 cccagctcgcttcagtggcagtgggtctgggacctcttactctctcacaatcagcagcat 239
          ||| |||||||||||||| ||||||||||||||||||||||||||||| |||||||  |
Qy    175 CCCTGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGT 234

Db    240 ggaggctgaagatgctgccacttattactgccaccagtatcatcgttccccgctcacgtt 299
          |||||||||||||||||||||||||||||| ||||    ||||  ||| | ||||||
Qy    235 GGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTT 294

Db    300 cggtgctgggaccaagctggagctcaaagtcgacaaatcctcaggatctggctccgaatc 359
          ||| |  ||| ||||||||||  | ||||    | |  |||| || |||||       ||
Qy    295 CGGAGGGGGGTCCAAGCTGGAAATAAAAGGTT-CTA--CCTCTGGTTCTGGTAAATCTTC 351

Db    360 caaaagcacgcaggtcaaactcgaggagtctggatctgagctggtgaggcctggagcttc 419
            || | |     ||  | ||  ||||||| ||| ||||| |||||| ||||||| |||
Qy    352 TGAAGGTAAAGGTGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTC 411

Db    420 agtgaagctgtcctgcaaggcttctggctacacattcaccagctactggatgcactgggt 479
          | ||||| | ||||||||| ||||||| ||| |||||||  || ||   ||| |||||||
Qy    412 AATGAAGATATCCTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGT 471

Db    480 gaagcagaggcatggacaaggccttgagtggattggaaatatttatcctggtagtggtag 539
          |||||||| |||||| |   ||||||||||||||||  |||| |||||   | ||||
Qy    472 GAAGCAGAGCCATGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGA 531

Db    540 tactaactacgatgagaagttcaagagcaaggccacactgactgtagacacatcctccag 599
          |||||||||| |  |||||||||| |||||||||| | ||||||||||   || |||||
Qy    532 TACTAACTACAACCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAG 591

Db    600 cacagcctacatgcacctcagcagcctgacatctgaggactctgcggtctattactgtac 659
          |||||||||||| | |||  ||| ||||||||||||||||||| |||||||||| |
Qy    592 CACAGCCTACATGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGC 651

Db    660 aagaagcagccggaactgggtctactatg 688
          ||| ||      | |||||  || | |||
Qy    652 AAGGAGGGTTACGGACTGGTACTTCGATG 680


RESULT    8
ID   Q73679 standard; DNA; 729 BP.
AC   Q73679;
DT   09-MAY-1995  (first entry)
DE   Fv(GP-4) immunosuppressive.
KW   MAb; monoclonal antibody; hybridoma; interleukin-2; IL-2;
KW   Fv; antibody variable region; GP-4; Fv(GP-4);
KW   immunosuppressive; ss.
OS   Mus sp.
PN   EP-621338-A.
PD   26-OCT-1994.
PF   21-APR-1994; 106257.
PR   21-APR-1993; JP-094491.
PR   07-MAR-1994; JP-036065.
```

PA (AJIN) AJINOMOTO KK.
PI Hamura J, Kanayama Y, Nakazawa H, Shimamura T, Sugamura K;
PI Takeshita T;
DR WPI; 94-325948/41.
DR P-PSDB; R60781.
PT Immunosuppressant polypeptide - has ability to block
PT interleukin-2 response
PS Claim 18; Page 29; 37pp; English.
CC MAb capable of binding to the gamma chain of the IL-2 receptor, and
CC thus of blocking the IL-2 response, is produced by mouse hybridoma
CC line GP-4 (FERM BP-4640). DNA encoding the variable region of
CC this MAb was expressed in E. coli, yielding Fv(GP-4) with
CC immumosuppressive activity.
SQ Sequence 729 BP; 199 A; 169 C; 182 G; 179 T;

Query Match 39.2%; Score 279; DB 12; Length 729;
Best Local Similarity 73.8%; Pred. No. 4.18e-177;
Matches 487; Conservative 0; Mismatches 166; Indels 7; Gaps 5;

```
Db      1 atggatattctgctgacacagtctccagcctccctatctgcatctgtgggagaaactgt- 59
          |||||  | | ||||||| ||||||||||||   ||| |||||||||   || || |  ||
Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db     60 caccatcacatgtcgagcaagtgggaatattcacaattatttagcatggtatcagcagaa 119
          ||| || || ||  | || |     | | |  | | |||  |    ||||||||||||||
Qy     61 CACAATGACTTGCAGGGCCACCCCAAGTGT--A-AGTTACATGCACTGGTATCAGCAGAA 117

Db    120 acagggaaaatctcctcagctcctggtctataatgcaaaaaccttagcagatggtgtgcc 179
            |  |||   || ||  | |    | | ||||  ||   | | | ||  ||| || ||
Qy    118 GCCAGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCC 177

Db    180 atcaaggttcagtggcagtggatcaggaacacaatattctctcaagatcaacagcctgca 239
            |  | |||||||||| |||| || || ||    || |||||||   ||| |||  || |
Qy    178 TGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGA 237

Db    240 gcctgaagattttgggagttattactgtcaacatttttggagtactccgtggacgttcgg 299
          | ||||||||  ||  | |||||||||| || || |   |  |||  ||    ||||||||
Qy    238 GGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGG 297

Db    300 tggagggaccaagctggagctcaaagtcgagaaatcctcaggatctggctccgaatccaa 359
            || ||| ||||||||||  | ||||  |    | |||| || |||||        ||  |
Qy    298 AGGGGGGTCCAAGCTGGAAATAAAAG--GTTCTA-CCTCTGGTTCTGGTAAATCTTCTGA 354

Db    360 aagcacgcaggtcaaactcgaggagtctggacctgagctggtgaagcctggggcttcagt 419
          | | |     ||  | ||  |||||||| |||||||| |||||||||||| | |||| |
Qy    355 AGGTAAAGGTGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAAT 414

Db    420 gaagatatcctgcaaggcttctggttactcattcactggctactacatgcactgggtgaa 479
          ||||||||||||||||  ||||||||||||||||||||||||  ||  |||| ||||||||||
Qy    415 GAAGATATCCTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAA 474

Db    480 gcaaagccatgtaaagagccttgagtggattggacgtattaatccttacaatggtgctac 539
          |||  ||||||| ||||| |||||||||||||||| ||||||||||||||||||||  |||
Qy    475 GCAGAGCCATGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATAC 534

Db    540 tagctacaaccagaatttcaaggacaaggccagcttgactgtagataagtcctccagcac 599
          ||  |||||||||||| |||||||  ||||||||  || ||||||||  |||||  ||||||||
Qy    535 TAACTACAACCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCAC 594

Db    600 agcctacatggagctccacagcctgacatctgaggactctgcagtctattactgtgcaag 659
          ||||||||||||||||  |||  ||||||||||||||||||||||||||||||||||||||
Qy    595 AGCCTACATGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654
```

RESULT 9
ID T31541 standard; cDNA; 462 BP.
AC T31541;
DT 18-SEP-1996 (first entry)
DE 3H1 heavy chain variable region cDNA.
KW Anti-idiotype; monoclonal antibody; MAb; 3H1; CEA;
KW carcinoembryonic antigen; cancer; gene therapy; immunotherapy;
KW vaccine; genetic immunisation; heavy chain; VH; ss.
OS Mus sp.
FH Key Location/Qualifiers
FT cds 22..462
FT /*tag= a
FT /product= 3H1 heavy chain variable region
PN WO9620277-A2.
PD 04-JUL-1996.
PF 28-DEC-1995; U17103.
PR 28-DEC-1994; US-365484.
PA (KENT) UNIV KENTUCKY.
PI Chatterjee M, Chatterjee SK, Foon KA, Kohler H;
DR WPI; 96-321850/32.
DR P-PSDB; R98411.
PT Recombinant monoclonal anti-idiotype antibody 3H1 sequences - used
PT to develop prods. for the detection and treatment of
PT carcinoembryonic antigen-associated diseases, partic. cancers
PS Claim 6; Fig 2A; 121pp; English.
CC A cDNA clone (T31541) codes for the heavy chain variable region
CC (VH) (R98411) of 3H1, a murine monoclonal anti-idiotype antibody
CC that mimics a specific epitope of the 180,000 mol.wt.
CC carcinoembryonic antigen (CEA) and which elicits an immune response
CC in patients with advanced CEA-associated disease, e.g. colorectal
CC cancer. It was obtd. by PCR amplification of 3H1 hybridoma DNA
CC (see also T31545-46). The isolated polynucleotide, and/or a
CC polynucleotide (see also T31540) coding for the VL region (R98410)
CC of 3H1 can be used to design probes and primers, in expression
CC systems, and in pharmaceutical applns., including vaccines, gene
CC therapy and genetic immunisation, partic. against cancer.
SQ Sequence 462 BP; 106 A; 124 C; 117 G; 115 T;

Query Match 37.4%; Score 266; DB 21; Length 462;
Best Local Similarity 90.0%; Pred. No. 1.13e-167;
Matches 316; Conservative 0; Mismatches 32; Indels 3; Gaps 2;

```
Db     82 gtccagctgcaacagtctggacctgagctggtgaagcctggagcttcactgaagatttcc 141
          || ||||||||  |||| |||||||||| |||||||||||||||| |||| ||||||| |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db    142 tgcgaggcttctggttactcactcactgcctacaccatgaactgggtgaagcagagccat 201
          ||| || |||||||||||||| |||||| | ||||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    202 ggaaagagccttgagtgggttgggctgattaatcctttcagtggtgatactaactacagc 261
          ||||||| |||||||||| |||| || |||||||||| || ||||||||||||||||| |
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    262 cagaaattcacgggcaaggccacattaactgtagacaggtcatccagcacagcctacatg 321
          ||||| |||| ||||||||||||||| |||||||||| ||| ||||||||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    322 gagctcctcagtctgacatctgaggactctgcagtctattactgtgtcattactccggtt 381
          |||||||||||||||||||||||||||||||||||||||||||||||| ||  |    ||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTG-CA--AGGAGGGTT 661

Db    382 ccctactggtacttcgatgtctggggcgcagggaccacggtcaccgtctcc 432
            |  |||||||||||||||||||||||||||||||||||||||||||||||
Qy    662 ACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTCC 712
```

RESULT 10
ID T99435 standard; cDNA; 462 BP.
AC T99435;
DT 27-APR-1998 (first entry)
DE Anti-idiotype antibody 3H1 heavy chain variable region cDNA.
KW Anti-idiotype antibody 3H1; carcinoembryonic antigen; CEA;
KW tumour-specific antigen; tumour; colorectal cancer; lung cancer;
KW adenocarcinoma; therapy; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT CDS 22..462
FT /*tag= a
PN WO9738725-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U05953.

```
PR   09-APR-1997; US-631085.
PR   12-APR-1996; US-631085.
PA   (KENT ) UNIV KENTUCKY.
PI   Chatterjee M, Chatterjee SK, Foon KA;
DR   WPI; 97-526218/48.
DR   P-PSDB; W26729.
PT   Use of anti-idiotype antibody 3H1 - for delaying development of
PT   carcinoembryonic antigen associated tumours
PS   Disclosure; Page 39; 66pp; English.
CC   This cDNA sequence includes a coding region for the heavy chain
CC   variable region (VH) (see W26729) of murine monoclonal anti-idiotype
CC   antibody 3H1.  A cDNA sequence (see T99434) for the VL region (see
CC   W26728) is also provided.  These clones can be utilised in the
CC   recombinant production of 3H1 polypeptides.  3H1 induces a specific
CC   immune response against a distinct and specific epitope of
CC   carcinoembryonic antigen (CEA), a tumour-associated antigen.  The
CC   invention provides methods of delaying development of CEA-associated
CC   tumours using 3H1, particularly in high-risk individuals.  The
CC   method can be used to delay the development of a variety of tumours,
CC   including colorectal, gastric and pancreatic tumours, as well as
CC   other adenocarcinomas such as those of the breast and lung, and
CC   biliary cancer, and gynaecological cancers.
SQ   Sequence   462 BP;    106 A;    124 C;    117 G;    115 T;

  Query Match           37.4%;  Score 266;  DB 38;  Length 462;
  Best Local Similarity  90.0%;  Pred. No. 1.13e-167;
  Matches   316;  Conservative    0;  Mismatches  32;  Indels   3;  Gaps   2;

Db     82 gtccagctgcaacagtctggacctgagctggtgaagcctggagcttcactgaagatttcc 141
           || ||||||||  |||| |||||||||| ||||||||||||||| |||| ||||||| |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db    142 tgcgaggcttctggttactcactcactgcctacaccatgaactgggtgaagcagagccat 201
           ||| || ||||||||||||||| |||||| | ||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    202 ggaaagagccttgagtgggttgggctgattaatcctttcagtggtgatactaactacagc 261
           ||||||| |||||||||| |||| || |||||||||| || |||||||||||||||| |
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    262 cagaaattcacgggcaaggccacattaactgtagacaggtcatccagcacagcctacatg 321
           ||||| |||| |||||||||||||| |||||||||| ||| ||||||||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    322 gagctcctcagtctgacatctgaggactctgcagtctattactgtgtcattactccggtt 381
           |||||||||||||||||||||||||||||||||||||||||||| ||  |    ||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTG-CA--AGGAGGGTT 661

Db    382 ccctactggtacttcgatgtctggggcgcagggaccacggtcaccgtctcc 432
            |  ||||||||||||||||||||||||||||||||||||||||||||||||
Qy    662 ACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTCC 712


RESULT   11
ID   T34542 standard; cDNA; 462 BP.
AC   T34542;
DT   11-OCT-1996  (first entry)
DE   Monoclonal anti-idiotype antibody 3H1 VH cDNA.
KW   Anti-idiotype antibody; monoclonal antibody; CEA;
KW   carcinoembryonic antigen; tumour-associated antigen; cancer;
KW   vaccine; immunotherapy; ss.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   cds             22..462
FT                   /*tag=  a
PN   WO9620219-A2.
PD   04-JUL-1996.
PF   28-DEC-1995; U17105.
PR   28-DEC-1994; US-365484.
PA   (KENT ) UNIV KENTUCKY.
PI   Chatterjee M,  Chatterjee SK,  Foon KA,  Kohler H;
DR   WPI; 96-321809/32.
DR   P-PSDB; R99687.
PT   Monoclonal anti-idiotype antibody 3H1 - elicits an immune response
PT   to carcinoembryonic antigen
PS   Example 2; Fig 2A; 102pp; English.
CC   A cDNA clone (T34542) codes for the heavy chain variable region
CC   (R99687) of monoclonal anti-idiotype antibody 3H1 (ATCC HB 12003).
CC   It was obtd. by PCR amplification (see also T34543-44) of cDNA
CC   derived from 3H1 hybridoma cells.  Anti-idiotype antibody 3H1
CC   elicits a specific immune response to a unique epitope of
CC   carcinoembryonic antigen (CEA) that is not present on other
CC   members of the CEA family or on normal adult tissues.  3H1
CC   can be used as a vaccine to elicit immune responses in patients
CC   with advanced CEA-associated disease or, when labeled, to enhance
CC   tumour detection in imaging.
SQ   Sequence   462 BP;    106 A;    124 C;    117 G;    115 T;

  Query Match           37.4%;  Score 266;  DB 22;  Length 462;
  Best Local Similarity  90.0%;  Pred. No. 1.13e-167;
  Matches   316;  Conservative    0;  Mismatches  32;  Indels   3;  Gaps   2;

Db     82 gtccagctgcaacagtctggacctgagctggtgaagcctggagcttcactgaagatttcc 141
           || ||||||||  |||| |||||||||| ||||||||||||||| |||| ||||||| |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db    142 tgcgaggcttctggttactcactcactgcctacaccatgaactgggtgaagcagagccat 201
           ||| || ||||||||||||||| |||||| | ||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    202 ggaaagagccttgagtgggttgggctgattaatcctttcagtggtgatactaactacagc 261
           ||||||| |||||||||| |||| || |||||||||| || |||||||||||||||| |
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    262 cagaaattcacgggcaaggccacattaactgtagacaggtcatccagcacagcctacatg 321
           ||||| |||| |||||||||||||| |||||||||| ||| ||||||||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    322 gagctcctcagtctgacatctgaggactctgcagtctattactgtgtcattactccggtt 381
           |||||||||||||||||||||||||||||||||||||||||||| ||  |    ||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTG-CA--AGGAGGGTT 661

Db    382 ccctactggtacttcgatgtctggggcgcagggaccacggtcaccgtctcc 432
            |  ||||||||||||||||||||||||||||||||||||||||||||||||
Qy    662 ACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTCC 712


RESULT   12
ID   Q65630 standard; DNA; 384 BP.
AC   Q65630;
DT   01-FEB-1995  (first entry)
DE   Murine variable region light chain from 2BS.
KW   B cell lymphoma chimeric antibody; CD20; peripheral blood cells;
KW   cell lysis; ss.
OS   Mus musculus.
PN   WO9411026-A.
PD   26-MAY-1994.
PF   12-NOV-1993; U10953.
PR   13-NOV-1992; US-978891.
PR   03-NOV-1993; US-149099.
PA   (IDEC-) IDEC PHARM CORP.
PI   Anderson DR,  Hanna N,  Leonard JE,  Newman RA,  Rastetter WH;
PI   Reffme;
DR   WPI; 94-183162/22.
DR   P-PSBD; R55214.
PT   Treating B cell lymphoma with chimeric antibody - against CD20,
PT   causing rapid depletion of peripheral B cells, also new
PT   antibodies and hybridomas
PS   Disclosure; Fig 4; 101pp; English.
CC   The sequence is the murine variable region light chain derived from
CC   murine anti-CD20 monoclonal antibody 2BS.
CC   See also Q65629-35.
SQ   Sequence   384 BP;    92 A;    106 C;    92 G;    94 T;

  Query Match           37.2%;  Score 265;  DB 11;  Length 384;
  Best Local Similarity  93.8%;  Pred. No. 5.98e-167;
```

```
  Matches   287;  Conservative    1;  Mismatches  17;  Indels   1;  Gaps   1;

Db     80 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 138
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db    139 gccagctcaagtgtaagttacatccactggttccagcagaagccaggatcctcccccaaa 198
          |||| | |||||||||||||||| |||||||  ||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    199 ccctggatttatgccacatccaacctggcttctggagtccctgttcgcttcagtggcagt 258
          || ||||||||| |||||||||||||:|||||||||||||||| |||||||||||||| ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    259 gggtctgggacttcttactctctcaccatcagcagagtggaggctgaagatgctgccact 318
          ||||||||||| ||||||||||||||  |||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    319 tattactgccagcagtggactagtaacccacccacgttcggaggggggaccaagctggaa 378
          |||||||||||||||||| | ||| ||||||||||||||||||||| ||||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAA 316

Db    379 atcaaa 384
          || |||
Qy    317 ATAAAA 322


RESULT   13
ID   Q65629 standard; DNA; 9208 BP.
AC   Q65629;
DT   01-FEB-1995  (first entry)
DE   Vector contg. TCAE 8 DNA.
KW   B cell lymphoma chimeric antibody; CD20; peripheral blood cells;
KW   cell lysis; ss.
OS   Synthetic.
PN   WO9411026-A.
PD   26-MAY-1994.
PF   12-NOV-1993; U10953.
PR   13-NOV-1992; US-978891.
PR   03-NOV-1993; US-149099.
PA   (IDEC-) IDEC PHARM CORP.
PI   Anderson DR,  Hanna N,  Leonard JE,  Newman RA,  Rastetter WH;
PI   Reffme;
DR   WPI; 94-183162/22.
PT   Treating B cell lymphoma with chimeric antibody - against CD20,
PT   causing rapid depletion of peripheral B cells, also new
PT   antibodies and hybridomas
PS   Disclosure; Fig 3; 101pp; English.
CC   The sequence shows a vector contg. TCAE8, a gene encoding a chimeric
CC   anti-CD20 antibody for treatment of B cell lymphomas.  TCAE8
CC   contains 4 transcriptional cassettes, human Ig light and heavy chain
CC   constant regions, dihydrofolate reductase, neomycin phosphotransferase
CC   and murine variable regions.  The vector can be used to produce
CC   antibodies which cause depletion of peripheral blood B cells,
CC   including those associated with lymphoma.  They mediate complement-
CC   dependent lysis and lyse target cells by antibody-dependent cellular
CC   cytotoxicity.
CC   See also Q65629-35.
SQ   Sequence   9208 BP;    2237 A;    2399 C;    2388 G;    2182 T;

  Query Match           37.1%;  Score 264;  DB 11;  Length 9208;
  Best Local Similarity  93.2%;  Pred. No. 3.17e-166;
  Matches   288;  Conservative    1;  Mismatches  19;  Indels   1;  Gaps   1;

Db   1057 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 1115
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db   1116 gccagctgaagtgtaagttacatccactggttccagcagaagccaggatcctcccccaaa 1175
          |||| |  |||||||||||||||| |||||||  ||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db   1176 ccctggatttatgccacatccaacctggcttctggagtccctgttcgcttcagtggcagt 1235
          || |||||||| |||||||||||||:|||||||||||||||| |||||||||||||| ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db   1236 gggtctgggacttcttactctctcaccatcagcagagtggaggctgaagatgctgccact 1295
          ||||||||||| ||||||||||||||  |||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db   1296 tattactgccagcagtggactagtaacccacccacgttcggaggggggaccaagctggaa 1355
          |||||||||||||||||| | ||| ||||||||||||||||||||| ||||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAA 316

Db   1356 atcaaacgt 1364
          || ||| ||
Qy    317 ATAAAAGGT 325


RESULT   14
ID   T79900 standard; cDNA; 318 BP.
AC   T79900;
DT   27-DEC-1997  (first entry)
DE   Anti-Factor IX MAb chimeric light chain cDNA.
KW   Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;
KW   chimeric antibody; antibody engineering; light chain; ss.
OS   Chimeric Mus musculus.
OS   Chimeric Homo sapiens.
OS   Chimeric synthetic.
PN   WO9726010-A1.
PD   24-JUL-1997.
PF   17-JAN-1997; U00759.
PR   24-OCT-1996; US-029119.
PR   17-JAN-1996; US-010108.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
PI   Blackburn MN,  Church WR,  Feuerstein GZ,  Gross MS;
PI   Nichols AJ,  Padlan EA, Patel AH,  Sylvester DR;
DR   WPI; 97-385117/35.
DR   P-PSDB; W24532.
PT   Inhibiting thrombosis with self-limiting antibody to coagulation
PT   factor - avoids uncontrolled bleeding by providing only partial
PT   inhibition
PS   Example 7; Page 128; 150pp; English.
CC   This cDNA sequence encodes a mouse-human chimeric antibody
CC   light chain (W24532) in which the variable region is derived
CC   from mouse anti-human factor IX monoclonal antibody BC2 cDNA (see
CC   T79899) and human sequences from the immunoglobulin RF-TS3'CL
CC   framework.  It was obtained by PCR amplification (see T79897-98)
CC   of BC2 cDNA and insertion of the PCR product into F9HZHC 1-3 cDNA
CC   (see T77374).  Claimed anti-Factor IX chimeric antibodies are
CC   useful in the treatment of thrombosis.
SQ   Sequence   318 BP;    79 A;    91 C;    78 G;    70 T;

  Query Match           36.9%;  Score 263;  DB 34;  Length 318;
  Best Local Similarity  93.5%;  Pred. No. 1.68e-165;
  Matches   286;  Conservative    1;  Mismatches  18;  Indels   1;  Gaps   1;

Db     14 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 72
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db     73 gccagctcaagtgtaaattacatgcactggtaccagcagaagccaggatcctcccccaaa 132
          |||| | |||||||| |||||||||||||| ||||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    133 ccctggatttatgccacatccaacctggcttctggagtccctgctcgcttcagtggcagt 192
          || |||||||| |||||||||||||:|||||||||||||||||||||||||||||| ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    193 gggtctgggacctcttactctctcacaatcagcagagtggaggctgaagatgctgccact 252
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    253 tattactgccagcagtggagtattaacccacggacgttcggtggaggcaccaagctggaa 312
          ||||||||||||||||||||  || |||||  |||||||| || ||  ||||||||||||
```

```
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAA 316

Db    313 atcaaa 318
          || |||
Qy    317 ATAAAA 322


RESULT   15
ID   T77377 standard; cDNA; 321 BP.
AC   T77377;
DT   26-DEC-1997  (first entry)
DE   Mouse anti-human Factor IX antibody BC2 light chain cDNA.
KW   Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;
KW   humanised antibody; antibody engineering; light chain; CDR;
KW   complementarity determining region; myocardial infarction;
KW   angina; atrial fibrillation; stroke; kidney damage;
KW   pulmonary embolism; deep vein thrombosis; coronary angioplasty;
KW   disseminated intravascular coagulation; artificial organ; sepsis;
KW   shunt; prosthesis; ss.
OS   Mus musculus.
PN   WO9726010-A1.
PD   24-JUL-1997.
PF   17-JAN-1997; U00759.
PR   24-OCT-1996; US-029119.
PR   17-JAN-1996; US-010108.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
PI   Blackburn MN,  Church WR,  Feuerstein GZ,  Gross MS;
PI   Nichols AJ,  Padlan EA, Patel AH,  Sylvester DR;
DR   WPI; 97-385117/35.
DR   P-PSDB; W24520.
PT   Inhibiting thrombosis with self-limiting antibody to coagulation
PT   factor - avoids uncontrolled bleeding by providing only partial
PT   inhibition
PS   Example 5; Page 64; 150pp; English.
CC   This cDNA sequence encodes the light chain variable region (see
CC   W24520) of mouse anti-human Factor IX monoclonal antibody BC2.
CC   Claimed humanised antibodies (see W24510-18) contain CDRs (see
CC   W24504-09) of BC2 heavy and light chains inserted into framework
CC   regions of selected human antibody sequences.  They have self-
CC   limiting neutralising activity, and are useful as anticoagulant
CC   agents in treatment of thrombosis associated with myocardial
CC   infarction, unstable angina, atrial fibrillation, stroke, renal
CC   damage, pulmonary embolism, deep vein thrombosis, percutaneous
CC   transluminal coronary angioplasty, disseminated intravascular
CC   coagulation, sepsis, or artificial organs, shunts or prostheses
CC   (claimed).  Also claimed are chimeric antibodies (see T79900), and
CC   Fab and F(ab')2 fragments.  The claimed antibodies do not cause
CC   uncontrolled bleeding (contrast heparin and warfarin) since they
CC   provide only partial inhibition of coagulation.
SQ   Sequence   321 BP;    78 A;    92 C;    79 G;    72 T;

  Query Match           36.9%;  Score 263;  DB 34;  Length 321;
  Best Local Similarity  93.5%;  Pred. No. 1.68e-165;
  Matches   286;  Conservative    1;  Mismatches  18;  Indels   1;  Gaps   1;

Db     14 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 72
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db     73 gccagctcaagtgtaaattacatgcactggtaccagcagaagccaggatcctcccccaaa 132
          |||| | ||||||||| |||||||||||||| ||||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    133 ccctggatttatgccacatccaacctggcttctggagtccctgctcgcttcagtggcagt 192
          || ||||||||| |||||||||||||:|||||||||||||||||||||||||||||||| ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    193 gggtctgggacctcttactctctcacaatcagcagagtggaggctgaagatgctgccact 252
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    253 tattactgccagcagtggagtattaacccacggacgttcggtggaggcaccaagctggaa 312
          |||||||||||||||||||||  || |||||  |||||||| || ||  ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAA 316

Db    313 atcaaa 318
          || |||
Qy    317 ATAAAA 322


Search completed: Sat Jan  9 12:38:18 1999
Job time : 122 secs.
```

```
**************************************************************************
 ___   ___   _____   ____   _____   ____   _   _
| |\ \ / /| |  __| |  _ \ |  ___| |  _ \ | | | |
| | \ V / | | |__  | |_| | | |__  | |_| | | | | |
| |  \_/  | |  __| |  _ / |  __|  |  _ /  | |_| |
| |       | | |___ | | \ \ | |___ | | \ \  |  _  |
|_|       |_| |____||_|  \_\|_____||_|  \_\ |_| |_|  (TM)
**************************************************************************

        Release 3.1A John F. Collins, Biocomputing Research Unit.
        Copyright (c) 1993-1998  University of Edinburgh, U.K.
              Distribution rights by Oxford Molecular Ltd

MPsrch_nn   n.a. -  n.a. database search, using Smith-Waterman algorithm

Run on:          Sat Jan  9 11:32:13 1999;  MasPar time 954.49 Seconds
                                            1335.831 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-704-178-1
Description:     (1-712) from US08704178.seq (1 of 4)
Perfect Score:   712
N.A. Sequence:        1 ATGGACCTGCAGCTGACCCA.........GGACCACGGTCACCGTCTCC 712
     Comp:              TACCTGGACGTCGACTGGGT.........CCTGGTGCCAGTGGCAGAGG

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        embl-est55
                   1:em_est1 2:em_gss1 3:em_gss2 4:em_gss3
Database:        genbank-est107
                   5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
                   10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
                   14:gb_est18 15:gb_est19 16:gb_est2 17:gb_est20
                   18:gb_est21 19:gb_est3 20:gb_est4 21:gb_est5 22:gb_est6
                   23:gb_est7 24:gb_est8 25:gb_est9 26:gb_gss1 27:gb_gss2
                   28:gb_gss3 29:gb_gss4

Statistics:      Mean 10.709;  Variance 1.855;  scale 5.772

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES

                    %
Result              Query
   No.  Score      Match Length DB  ID          Description            Pred. No.
 ------------------------------------------------------------------------------
     1    232       32.6    396 11  AA691311    vs14f01.r1 Barstead mo  0.00e+00
     2    179       25.1    344  9  AA592800    vo25g11.r1 Barstead mo  0.00e+00
     3    160       22.5    310 18  AI180569    uc70f09.r1 Soares mous  3.63e-298
     4    146       20.5    294 15  AI007196    ua73g04.r1 Soares 2NbM  1.74e-265
     5    110       15.4    387  9  AA569186    nm30d10.s1 NCI_CGAP_Li  1.05e-182
     6     92       12.9    345 11  AA710291    vt53a04.r1 Barstead mo  2.72e-142
     7     86       12.1    335 24  AA318377    EST20620 Spleen I Homo  4.99e-129
     8     85       11.9    279 24  AA300571    EST13661 Testis tumor   7.88e-127
     9     82       11.5    379  8  AA291381    zt44g02.r1 Soares ovar  2.95e-120
    10     82       11.5    393 16  T27593      EST100653 Homo sapiens  2.95e-120
c   11     81       11.4    209 16  T28938      EST61186 Homo sapiens   4.51e-118
    12     81       11.4    488 16  H25625      yl48g05.r1 Homo sapien  4.51e-118
    13     80       11.2    333 24  AA300732    EST13847 Testis tumor   6.81e-116
    14     80       11.2    413 24  AA301347    EST14279 Testis tumor   6.81e-116
    15     79       11.1    116 17  AI116236    uc18e08.r1 Soares mous  1.02e-113
    16     79       11.1    194 14  AA921173    vy54b07.r1 Stratagene   1.02e-113
    17     79       11.1    253 24  AA295941    EST101165 Thymus III H  1.02e-113
    18     78       11.0    370 24  AA295093    EST100400 Pancreas tum  1.51e-111
    19     77       10.8    210 16  R69482      yj83c03.r1 Homo sapien  2.22e-109
    20     77       10.8    328 24  AA295311    EST100471 Pancreas tum  2.22e-109
    21     77       10.8    403 25  AA345486    EST51505 Gall bladder   2.22e-109
    22     76       10.7    242  8  AA479857    zu35b05.r1 Soares ovar  3.23e-107
    23     76       10.7    276 13  AA869306    vq50c10.r1 Barstead bo  3.23e-107
    24     76       10.7    431 11  AA710970    vt93h09.r1 Soares mous  3.23e-107
    25     75       10.5    352 24  AA301261    EST14181 Testis tumor   4.65e-105
    26     75       10.5    470 19  H62115      yu40h01.r1 Homo sapien  4.65e-105
    27     74       10.4    463 16  R69532      yj82d09.r1 Homo sapien  6.62e-103
    28     73       10.3    345 25  AA335086    EST39457 Esophagus tum  9.32e-101
    29     73       10.3    442 10  AA405415    zu56d02.r1 Soares ovar  9.32e-101
    30     71       10.0    287 16  T27609      EST101034 Homo sapiens  1.78e-96
    31     71       10.0    328 24  AA318628    EST20808 Spleen I Homo  1.78e-96
    32     70        9.8    503  5  R28232      yh57e08.r1 Homo sapien  2.42e-94
    33     69        9.7    339 11  AA710249    vt49e01.r1 Barstead mo  3.26e-92
    34     69        9.7    395 16  T29112      EST69384 Homo sapiens   3.26e-92
    35     69        9.7    438 16  R64693      yi22f09.r1 Homo sapien  3.26e-92
    36     69        9.7    467 16  H27034      yl65d05.r1 Homo sapien  3.26e-92
    37     68        9.6    287  8  AA496610    zv38g06.r1 Soares ovar  4.32e-90
    38     68        9.6    336 15  AA987559    or83g01.s1 NCI_CGAP_Lu  4.32e-90
    39     67        9.4    398 10  AA423447    ve80a03.r1 Soares mous  5.65e-88
    40     66        9.3    265 25  AA377128    EST89660 Small intesti  7.29e-86
    41     66        9.3    291 25  AA377074    EST89603 Small intesti  7.29e-86
    42     66        9.3    303 24  AA300891    EST14031 Testis tumor   7.29e-86
    43     66        9.3    363  7  AA464313    zx78c12.r1 Soares ovar  7.29e-86
    44     66        9.3    382 24  AA295786    EST100987 Pancreas tum  7.29e-86
    45     66        9.3    516 16  R67559      yi42h11.r1 Homo sapien  7.29e-86

                               ALIGNMENTS

RESULT    1
LOCUS       AA691311       396 bp    mRNA            EST       16-DEC-1997
DEFINITION  vs14f01.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone 1138201 5' similar to gb:X67211 M.musculus rearranged
            immunoglobulin kappa light chain (MOUSE);, mRNA sequence.
ACCESSION   AA691311
NID         g2692247
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 396)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:619473
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES              Location/Qualifiers
     source           1. .396
                      /organism="Mus musculus"
                      /strain="FVB/N"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
```

```
                polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                from 8 week old mouse.  Colon was harvested 72 hours after
                irradiation with 1400 Gys.  1st strand cDNA was primed
                with  a Not I - oligo(dT) primer
                [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
                T 3'];  double-stranded cDNA was ligated to Eco RI
                adaptors [AATTCGGATCCTTG],  digested with Not I and cloned
                into the Not I and Eco RI sites of  the modified pT7T3
                vector. Library constructed by Bob Barstead."
                /db_xref="taxon:10090"
                /clone="1138201"
                /clone_lib="Barstead mouse irradiated colon MPLRB7"
                /dev_stage="8 weeks"
                /lab_host="DH10B"
BASE COUNT      104 a    107 c     91 g     94 t
ORIGIN

  Query Match           32.6%;  Score 232;  DB 11;  Length 396;
  Best Local Similarity  89.6%;  Pred. No. 0.00e+00;
  Matches   266;  Conservative    1;  Mismatches  29;  Indels   1;  Gaps   1;

Db    101 CTCACCCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAGAAGGT-CACCATGACCTG 159
          || |||||||||||||||| |||||||||||||||||||||||||| ||| ||||| ||
Qy     13 CTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTG 72

Db    160 CAGTGCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGTCAGGCACCTCCCC 219
          ||| |||| | |||||||||||||||||||||||| |||||||| |||| |||||||
Qy     73 CAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCC 132

Db    220 CAAAAGATGGATTTATGACACATCCAAACTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGG 279
          ||||   |||||||||  |||||||||  ||:|||||||||||||||||||||||||||||
Qy    133 CAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGG 192

Db    280 CAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGCATGGAGGCTGAAGATGCTGC 339
          | |||||||||||||||||||||||||||||| |||||||  ||||||||||||||||||
Qy    193 CGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGC 252

Db    340 CACTTATTACTGCCAGCAGTGGAGTAGTAACCCGCTCACGTTCGGTGCTGGGACCAA 396
          |||||||||||||||||||||||| ||| ||| | |||||||| |  ||| ||||
Qy    253 CACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAA 309


RESULT    2
LOCUS       AA592800      344 bp    mRNA            EST       17-SEP-1997
DEFINITION  vo25g11.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone 1050980 5' similar to gb:X67211 M.musculus rearranged
            immunoglobulin kappa light chain (MOUSE);, mRNA sequence.
ACCESSION   AA592800
NID         g2406463
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 344)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:582556
            Trace considered overall poor quality
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .344
                      /organism="Mus musculus"
                      /strain="FVB/N"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                      from 8 week old mouse.  Colon was harvested 72 hours after
                      irradiation with 1400 Gys.  1st strand cDNA was primed
                      with  a Not I - oligo(dT) primer
                      [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
                      T 3'];  double-stranded cDNA was ligated to Eco RI
                      adaptors [AATTCGGATCCTTG],  digested with Not I and cloned
                      into the Not I and Eco RI sites of  the modified pT7T3
                      vector. Library constructed by Bob Barstead."
                      /db_xref="taxon:10090"
                      /clone="1050980"
                      /clone_lib="Barstead mouse irradiated colon MPLRB7"
                      /dev_stage="8 weeks"
                      /lab_host="DH10B"
     mRNA             <1. .>344
BASE COUNT      91 a     91 c     70 g     92 t
ORIGIN

  Query Match           25.1%;  Score 179;  DB 9;  Length 344;
  Best Local Similarity  83.7%;  Pred. No. 0.00e+00;
  Matches   226;  Conservative    1;  Mismatches  42;  Indels   1;  Gaps   1;

Db     75 CTCACTCAATCTCCAGCAATCATGTCTGCTTCTCCAGAGGGAGCAATGTCACCACGACCT 134
          || || || |||||||||||| ||||||| ||||||| ||||| |    ||| | ||| |
Qy     13 CTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAG-GGGAGAAGGTACACAATGACTT 71

Db    135 GCAGTGCCAGTTCTACTGTAAGTTACATGCAATGGTAACAGCAGAAGTCAGGTACCTCCC 194
          |||| ||||   | | |||||||||||||| ||||| ||||||||| |||| ||||||
Qy     72 GCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCC 131

Db    195 CCAAAACATGGATTTATGATACATCTAAACTGGCTTCTGGAATCCCTGCTCGCTTCAGTG 254
          ||||| | |||||||||   ||||| || ||:||||||||| ||||||||||||||||||
Qy    132 CCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTG 191

Db    255 GCAGTAGCTCTGGGACCTCCTACTCTCTCACAATCAGCAGAATTGAGGCTGAACATGCTG 314
          || || | |||||||||| |||||||||| |||||||| | ||||||||| ||||||
Qy    192 GCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTG 251

Db    315 CCACTTATTACTGCCAGCATTGGAGTAGTA 344
          |||||||||||||||||| |||||| |||
Qy    252 CCACTTATTACTGCCAGCAGTGGAGTCGTA 281


RESULT    3
LOCUS       AI180569      310 bp    mRNA            EST       08-OCT-1998
DEFINITION  uc70f09.r1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone
            1431017 5' similar to gb:M18512 IG HEAVY CHAIN PRECURSOR V-I REGION
            (HUMAN); gb:M12376 Mouse immunoglobulin H-chain V-region pseudogene
            mRNA, complete (MOUSE);, mRNA sequence.
ACCESSION   AI180569
NID         g3731207
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 310)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
```

```
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:915085
            Seq primer: -28ml3 rev2 ET from Amersham.
FEATURES              Location/Qualifiers
     source           1. .310
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Organ: mammary gland; Vector: pT7T3D-Pac
                      (Pharmacia) with a modified polylinker; Site_1: Not I;
                      Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                      oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTTTT
                      T 3']; double-stranded cDNA was ligated to Eco RI
                      adaptors (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of the modified pT7T3 vector.
                      RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                      constructed and normalized by Bento Soares and M.Fatima
                      Bonaldo."
                      /db_xref="taxon:10090"
                      /clone="1431017"
                      /clone_lib="Soares mouse mammary gland NbMMG"
                      /sex="male"
                      /tissue_type="mammary gland"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
BASE COUNT       80 a     77 c     85 g     67 t      1 others
ORIGIN

  Query Match           22.5%;  Score 160;  DB 18;  Length 310;
  Best Local Similarity  79.3%;  Pred. No. 3.63e-298;
  Matches   230;  Conservative    0;  Mismatches  58;  Indels   2;  Gaps   2;

Db       4 GTCCAACTGCAGCAGCCTGGGG-TAAGCTTGTGAAGCCTGG-GGTTCAGTGAAGCTGTCC 61
           || || |||||| || | ||   | || | |||||||||| |||||| ||||| | |||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db      62 TGCAAGGCTTCTGGCTACACCTTCACCAGCTACTGGATGCAGTGGGTAAAACAGAGGCCT 121
           |||||| ||||||| ||| | |||||  || ||   ||| | ||||| || ||||| | |
Qy     425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db     122 GGACAGGGCATTGAGTGGATCGGAGAGATTGATCCTTCTGATAGCTATACTAACTACAAT 181
           ||| ||  | |||||||||| ||| |||| |||||| || | ||||||||||||
Qy     485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db     182 CAAAAGTTCAAGGGCAAGGCAGCATTGACTGTAGACACATCCTCCAGCACAGCCTACATG 241
           || |||||||||||||||  |||| |||||||||| || ||||||||||||||||||
Qy     545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db     242 CAGCTCAGCAGCCTGACATCTGAGGACTCTGCGGTCTATGACTGTGCAAG 291
            |||||  ||| ||||||||||||||||||||| |||||| ||||||||||
Qy     605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654


RESULT    4
LOCUS       AI007196      294 bp    mRNA            EST       12-JUN-1998
DEFINITION  ua73g04.r1 Soares 2NbMT Mus musculus cDNA clone 1363158 5' similar
            to gb:M18512 IG HEAVY CHAIN PRECURSOR V-I REGION (HUMAN); gb:M12376
            Mouse immunoglobulin H-chain V-region pseudogene mRNA, complete
            (MOUSE);, mRNA sequence.
ACCESSION   AI007196
NID         g3216753
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 294)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:896378
            Trace considered overall poor quality
            Seq primer: -28ml3 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .294
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. RNA
                      provided by Dr. Bertrand Jordan. Library went through two
                      rounds of normalization, and was constructed by Bento
                      Soares and M.Fatima Bonaldo."
                      /db_xref="taxon:10090"
                      /clone="1363158"
                      /clone_lib="Soares 2NbMT"
                      /sex="male"
                      /tissue_type="Thymus"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
BASE COUNT       85 a     68 c     77 g     64 t
ORIGIN

  Query Match           20.5%;  Score 146;  DB 15;  Length 294;
  Best Local Similarity  76.7%;  Pred. No. 1.74e-265;
  Matches   224;  Conservative    0;  Mismatches  66;  Indels   2;  Gaps   2;

Db       2 GTCCAACTGCAGCAGCCTGGGGCTGAGCATGGTGAGGACTGGGTCTTCAGTGAAGCTGTA 61
           || || |||||| || | ||  |||||  |||||| | ||||   |||| ||||| | |
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGG-TGGTGAAGCCTGGAGGTTCAATGAAGATATC 423

Db      62 CTGCAAGGATTCAGGCTACACCTTCACCAGCTACTGGATGCATTGGGTGAAGCAGAGGCA 121
           |||||||  ||| || ||| | |||||  || ||   ||| | ||||||||||||||| ||
Qy     424 CTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCA 483

Db     122 TATACAAGGCATTGAATGGATTGGTAACATTGACCCTTCAGATAGTGGAACTCACTACAA 181
           |  | |   | |||| ||||||||    ||| | ||||   || |||  ||| |||||||
Qy     484 TGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAA 543

Db     182 TCAAAAGTTCAAGGACAAGGCCACATTGACTGTAGACAAATTCACCAGCACAGCCTACAG 241
            || |||||||||| |||||||||||| |||||||||||  |   ||||||||||||||
Qy     544 CCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACA- 602

Db     242 TGCAGCTCAGCAGACTGACATCTGAGGACTCTGCGGTCTATTACTGTGCAAG 293
           || |||||  ||| ||||||||||||||||||| ||||||||||||||||||
Qy     603 TGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654


RESULT    5
```

```
LOCUS       AA569186      387 bp    mRNA            EST       09-SEP-1997
DEFINITION  nm30d10.s1 NCI_CGAP_Lip2 Homo sapiens cDNA clone IMAGE:1061683
            similar to gb:L02325 IG HEAVY CHAIN PRECURSOR V-I REGION (HUMAN);,
            mRNA sequence.
ACCESSION   AA569186
NID         g2342240
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 387)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Insert Length: 1722   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 284.
FEATURES             Location/Qualifiers
     source          1. .387
                     /organism="Homo sapiens"
                     /note="Vector: pAMP10; mRNA made from liposarcoma, cDNA
                     made by oligo-dT priming. Non- directionally cloned.
                     Size-selected on agarose gel, average insert  size 600 bp.
                     Reference:  Krizman et al. (1996) Cancer Research
                     56:5380-5383."
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1061683"
                     /clone_lib="NCI_CGAP_Lip2"
                     /tissue_type="liposarcoma"
                     /lab_host="DH10B"
     mRNA            <1. .>387
BASE COUNT       89 a     95 c    118 g     85 t
ORIGIN

  Query Match           15.4%;  Score 110;  DB 9;  Length 387;
  Best Local Similarity  69.0%;  Pred. No. 1.05e-182;
  Matches   200;  Conservative    0;  Mismatches  90;  Indels   0;  Gaps   0;

Db      98 GTCCAACTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGGGGCCTCAGTGAAGCTTTCC 157
           || || |||  | |||| ||  |||||||| |||||||| |  ||| ||||| | |||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     158 TGTAAGGCATCTGGATTCACCCTCACCAACTGCCATATGCACTGGGTGCGACAGGCCCCT 217
           || ||| | ||||| | | |  ||||   |  |   ||| |||||||   |||  || |
Qy     425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db     218 GGGCAAGGGCTTGAGTGGGTGGGAATGATCAATTCTAGTGATGGTTATATAAGTAACGCA 277
           ||   |    ||||||||| | ||| | || ||| ||    ||||| |||  |   ||
Qy     485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db     278 CAGAAGTTCCAGGGCAGAGTCACCATGACCAGGGACACGTCCACGAGCACAGTCTATATG 337
           |||||||| ||||||  | |||  | ||    |||| ||| | ||||||| ||| |||
Qy     545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db     338 GAGCTTAGCAGCCTGAGATCTGAGGACACGGCCGTCTATTACTGTGCGAG 387
           |||||   ||| |||| |||||||||| | || ||||||||||||| ||
Qy     605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654
```

```
RESULT    6
LOCUS       AA710291      345 bp    mRNA            EST       24-DEC-1997
DEFINITION  vt53a04.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone 1166766 5' similar to gb:X02484_rna1 IG KAPPA CHAIN PRECURSOR
            V-IV REGION (HUMAN); gb:J00560 mouse ig kappa mrna from mopc21 &
            other myeloma mrna 3' (MOUSE);, mRNA sequence.
ACCESSION   AA710291
NID         g2720209
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 345)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:632678
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 140.
FEATURES             Location/Qualifiers
     source          1. .345
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                     from 8 week old mouse.  Colon was harvested 72 hours after
                     irradiation with 1400 Gys.  1st strand cDNA was primed
                     with  a Not I - oligo(dT) primer
                     [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTT
                     T 3'];  double-stranded cDNA was ligated to Eco RI
                     adaptors [AATTCGGATCCTTG],  digested with Not I and cloned
                     into the Not I and Eco RI sites of  the modified pT7T3
                     vector. Library constructed by Bob Barstead."
                     /db_xref="taxon:10090"
                     /clone="1166766"
                     /clone_lib="Barstead mouse irradiated colon MPLRB7"
                     /dev_stage="8 weeks"
                     /lab_host="DH10B"
BASE COUNT       95 a     81 c     88 g     81 t
ORIGIN

  Query Match           12.9%;  Score 92;  DB 11;  Length 345;
  Best Local Similarity  66.9%;  Pred. No. 2.72e-142;
  Matches   208;  Conservative    1;  Mismatches  99;  Indels   3;  Gaps   3;

Db      35 TGACCCAGTCTCAAAAATTCATGTCCACATCAGCTGGAGACAGGGT-CAGCGTCACCTGC 93
           |||||||||||| |  | || ||||  ||||  | || || | ||| ||   | || |||
Qy      14 TGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGC 73

Db      94 AAGCCAGTCAGAATGTGGGTACTAATGTAGCCTGGTATCAAGAGAAACCAGGGCAATCTC 153
           | | |   |  || |||   | | |  |   |||||||||  |||| |||||    || |
Qy      74 AGGGCCACCCCAA-GTGTA-AGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCC 131

Db     154 CTAGAGCACTGATTTACTCGGCATCCTACCGGTACAGTGGAGTCCCTGATCGCTTCACAG 213
           | | | |   ||||||  |  ||||| ||| :     |||||||||||| ||||||||  |
Qy     132 CCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTG 191
```

```
Db    214 GCAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAATGTGCAGTCTGAAGACTTGG 273
          || |||| ||||||||   || | |||||||  ||||||  ||| || |||||||    |
Qy    192 GCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTG 251

Db    274 CAGAGTATTTCTGTCAGCAATATAACAGCTTTCCGTACACGTTCGGAGGGGGGACCAAGC 333
          |    |||| ||| ||||| |  |   |    ||   ||||||||||||||||| ||||||
Qy    252 CCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGC 311

Db    334 TGGAAATAAAA 344
          |||||||||||
Qy    312 TGGAAATAAAA 322


RESULT    7
LOCUS       AA318377      335 bp    mRNA            EST       19-APR-1997
DEFINITION  EST20620 Spleen I Homo sapiens cDNA 5' end similar to similar to
            immunoglobulin kappa light chain, V region, anti-thyroglobulin
            (GB:X79786), mRNA sequence.
ACCESSION   AA318377
NID         g1970863
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 335)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES              Location/Qualifiers
     source           1. .335
                      /organism="Homo sapiens"
                      /note="Organ: spleen; Vector: pBluescript SK-; Site_1:
                      EcoRI; Site_2: XhoI"
                      /db_xref="ATCC (inhost):119058"
                      /db_xref="taxon:9606"
                      /clone_lib="Spleen I"
                      /sex="male"
                      /dev_stage="adult, 23 yrs"
     mRNA             <1. .>335
BASE COUNT       87 a     88 c     71 g     84 t      5 others
ORIGIN
```

```
  Query Match           12.1%;  Score 86;  DB 24;  Length 335;
  Best Local Similarity  67.1%;  Pred. No. 4.99e-129;
  Matches   159;  Conservative    1;  Mismatches  77;  Indels   0;  Gaps   0;

Db     36 AGCTATTTACATTGGTATCAGCACAAACCAGGTAAAGCCCCTAAACTCCTGATCTATGNT 95
          || ||  | || |||||||||||| || |||||     |||| ||||    ||| |||
Qy     92 AGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAACCTTGGATTTATACC 151

Db     96 ACATCCACTTTGCAAAGTGGGGTCCCATCAAGGTTCAGTGGCACTGGATCTGGGACAGAT 155
          |||||||   |:     ||| |||||  |  | |||||||||  ||| ||||||||   |
Qy    152 ACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCT 211

Db    156 TTCACTCTCACCATCAGCAGTCTGCAACCTGAAGATTTTGCAACTTACTACTGTCAACAG 215
          | | |||||||  |||||||  || |  ||||||||  ||| |||||  ||||| || |||
Qy    212 TACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAG 271

Db    216 AGTTCCAGTAGCCCTTGGACGTTCGGCCAAGGGNCCAAGGTGGCAATCAAACGNACT 272
          |     |||||||    ||||||||    ||| ||||| ||| ||| ||| |  ||
Qy    272 TGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAAATAAAAGGTTCT 328


RESULT    8
LOCUS       AA300571      279 bp    mRNA            EST       18-APR-1997
DEFINITION  EST13661 Testis tumor Homo sapiens cDNA 5' end similar to similar
            to immunoglobulin mu heavy chain, VDJC regions (GB:M18512), mRNA
            sequence.
ACCESSION   AA300571
NID         g1953132
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 279)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: THC166571
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES              Location/Qualifiers
     source           1. .279
                      /organism="Homo sapiens"
                      /note="Organ: testis; Vector: pBluescript SK-; Site_1:
```

```
                     EcoRI; Site_2: XhoI"
                     /db_xref="ATCC (inhost):192212"
                     /db_xref="taxon:9606"
                     /clone_lib="Testis tumor"
                     /sex="male"
                     /dev_stage="adult"
     mRNA            <1. .>279
BASE COUNT       63 a     74 c     85 g     56 t      1 others
ORIGIN

  Query Match           11.9%;  Score 85;  DB 24;  Length 279;
  Best Local Similarity  72.3%;  Pred. No. 7.88e-127;
  Matches   136;  Conservative    0;  Mismatches  52;  Indels   0;  Gaps   0;

Db      92 GTGCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGGGGCCTCAGTGAAGGTCTCC 151
           ||||||||||  | ||||  ||  ||||||||  |||||||||  |  ||| |||||  | |||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     152 TGCAAGGCTTCTGGATACACCTTCACCGGCTACTATATGCACTGGGTGCGACAGGCCCCT 211
           ||||||  |||||||  |||  |  |||||  |||  ||    |||  ||||||||    |||   || |
Qy     425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db     212 GGACAAGGGCTTNAGTGGATGGGATGGATCAACCCTAAAAGTGGTGGCACAAACTATGCA 271
           |||  |     |||  |||||||  |||    ||  ||  |||  |  |  |||||   ||  |||||
Qy     485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db     272 CAGAAGTT 279
           ||||||||
Qy     545 CAGAAGTT 552


RESULT    9
LOCUS       AA291381      379 bp    mRNA            EST       08-AUG-1997
DEFINITION  zt44g02.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 725234
            5' similar to gb:M18512 IG HEAVY CHAIN PRECURSOR V-I REGION
            (HUMAN);, mRNA sequence.
ACCESSION   AA291381
NID         g1939359
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 379)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1576   Std Error: 0.00
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 361.
FEATURES              Location/Qualifiers
     source           1. .379
                      /organism="Homo sapiens"
                      /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library constructed by Bento Soares and
                      M.Fatima Bonaldo."
                      /db_xref="taxon:9606"
                      /clone="725234"
                      /clone_lib="Soares ovary tumor NbHOT"
                      /sex="Female"
                      /tissue_type="ovarian tumor"
                      /lab_host="DH10B (ampicillin resistant)"
     mRNA             <1. .>379
BASE COUNT       85 a     82 c    121 g     91 t
ORIGIN

  Query Match           11.5%;  Score 82;  DB 8;  Length 379;
  Best Local Similarity  65.9%;  Pred. No. 2.95e-120;
  Matches   191;  Conservative    0;  Mismatches  97;  Indels   2;  Gaps   2;

Db      37 GTTCAGTTGGTGCAGTCTGGACCTGAGGTGAAGAAGC-TGG-GGCCCAGTGAGGGTCTCC 94
           || ||| ||  | |||| ||||||||||||  ||||| ||| ||  || ||| | | |||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db      95 TGTAAGACCTCTGGTTACACCTTTGGCACTTTTGGTATCAACTGGGTGCGGCAGGCCCCT 154
           || ||||| |||||||||  | ||              || ||||||||  ||||  || |
Qy     425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db     155 GGACAAGGGCTTGAGTGGATGGGATGGATCAGCGTTTACAATGGTAAGACAAACTTTGCA 214
           ||| |    ||||||||||| |||   || |    |||||||||| | || ||||
Qy     485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db     215 GGGAGTTTTCAGGACAGAGTCAGCTTGACCACTGACACATCCACGAATACAGCCTACATG 274
           ||  ||  ||| ||  | ||  || ||    ||||  ||  | |  ||||||||||||
Qy     545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db     275 GAGCTGAGGAACCTGAGATCTGACGACACGGCCATATATTACTGTGCGAG 324
           |||||    |  |||| ||||||  ||| | ||  | ||||||||||| ||
Qy     605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654


RESULT   10
LOCUS       T27593        393 bp    mRNA            EST       06-SEP-1995
DEFINITION  EST100653 Homo sapiens cDNA 5' end similar to immunoglobulin kappa
            light chain, V region (GB:L01279) (HT:3043).
ACCESSION   T27593
NID         g609691
KEYWORDS    EST.
SOURCE      human primer=M13 Reverse library=Human Pancreas.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 393)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W.,
            Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            FitzGerald,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghagen,N.S.M.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle Jr,P.S.,
            Kelley,J.M., Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M.,
            Merrick,J.M., Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T.,
            Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L.,
            Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R.,
            Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A.,
            Coleman,T.A., Collins,E.-J., Dimke,D., Feng,P., Ferrie,A.,
            Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M.,
            Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H.,
            Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C.,
            Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A.,
            Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
  TITLE     Initial Assessment of Human Gene Diversity and Expression Patterns
            Based Upon 52 Million Basepairs of cDNA Sequence
  JOURNAL   Unpublished (1995)
COMMENT     Other_ESTs: THC24356
            Contact: Venter, JC
```

```
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: tdbinfo@tdb.tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please contact the TIGR Database
            (tdbinfo@tdb.tigr.org).
FEATURES              Location/Qualifiers
     source           1. .393
                      /organism="Homo sapiens"
     mRNA             <1. .>393
BASE COUNT       88 a    113 c     96 g     94 t      2 others
ORIGIN

  Query Match           11.5%;  Score 82;  DB 16;  Length 393;
  Best Local Similarity  65.9%;  Pred. No. 2.95e-120;
  Matches   213;  Conservative    1;  Mismatches 104;  Indels   5;  Gaps   3;

Db      53 GACATCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACAGAGT-CAC 111
           ||| | ||| ||||||||||||||| |   |||||||||||||  ||| || |  || |||
Qy       4 GACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACAC 63

Db     112 CCTCACTTGCCGGGCAAGTCAGAGCATTAGCACCTTTTTAAATTGGTATCAGCAGAAACC 171
           | |||||| |||| |  |  ||  | ||   | |     | ||||||||||||||| ||
Qy      64 AATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGC---ACTGGTATCAGCAGAAGCC 120

Db     172 AGGGACAGCCCCTAAGCTCCTGATCTCTGCTGCATCCAATTTGCGAAGTGGGGTCCCATC 231
           |||  |  |||| || |    ||| | | |  ||||||  |:     ||| |||||  |
Qy     121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180

Db     232 AAGGTTCAGTGGCTCTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGTCTGCAACC 291
           | |||||||||  ||| |||||||   || | |||||||  |||||||  || |  |
Qy     181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db     292 TGAAGATGTTGCAACTTACTACTGTCAACAGAGTTNCACTTATGNCTCGGACGGTTGGCC 351
           |||||||| ||| ||||| ||||| || ||| |     | ||   | |  ||| | ||
Qy     241 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGT-CGTAGCCCACCCACGTTCGGAG 299

Db     352 AAGGGACCAAGGTGGAGGTCAAA 374
           ||| ||||| ||||  | |||
Qy     300 GGGGGTCCAAGCTGGAAATAAAA 322


RESULT   11
LOCUS       T28938        209 bp    mRNA            EST       06-SEP-1995
DEFINITION  EST61186 Homo sapiens cDNA 5' end similar to immunoglobulin mu
            heavy chain, VDJ regions (GB:M17751) (HT:3055).
ACCESSION   T28938
NID         g611036
KEYWORDS    EST.
SOURCE      human primer=M13 Reverse library=Human White blood cells.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 209)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W.,
            Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            FitzGerald,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghagen,N.S.M.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle Jr,P.S.,
            Kelley,J.M., Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M.,
            Merrick,J.M., Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T.,
            Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L.,
            Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R.,
            Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A.,
            Coleman,T.A., Collins,E.-J., Dimke,D., Feng,P., Ferrie,A.,
            Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M.,
            Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H.,
            Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C.,
            Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A.,
            Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
  TITLE     Initial Assessment of Human Gene Diversity and Expression Patterns
            Based Upon 52 Million Basepairs of cDNA Sequence
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Venter, JC
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: tdbinfo@tdb.tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please contact the TIGR Database
            (tdbinfo@tdb.tigr.org).
FEATURES              Location/Qualifiers
     source           1. .209
                      /organism="Homo sapiens"
     mRNA             <1. .>209
BASE COUNT       43 a     62 c     48 g     55 t      1 others
ORIGIN

  Query Match           11.4%;  Score 81;  DB 16;  Length 209;
  Best Local Similarity  71.7%;  Pred. No. 4.51e-118;
  Matches   134;  Conservative    0;  Mismatches  53;  Indels   0;  Gaps   0;

Db      15 GTAATGGTGACTCTGCCCTGGAACTTCTGTAAATATTTTGTGTTACCATTGCCAGCGTTG 74
           ||||  ||| |  |||||| |||||||||    || || || | |||||||  ||  ||
Cp     573 GTAAATGTGGCCTTGCCCTTGAACTTCTGGTTGTAGTTAGTATCACCATTGTAAGGATTA 514

Db      75 ATCCATCCCATCCACTCAAGCCTTTGTCCGGGGGCCTGGCGCACCCAATGCATAGCATAG 134
           ||   ||| ||||||||||  | | |||  ||  |||   |||||| | ||| |  | |
Cp     513 ATAAGTCCAATCCACTCAAGGTTCTTTCCATGGCTCTGCTTCACCCAGTTCATGGTGTGG 454

Db     135 CTAGTGAAGGTGTATCCAGAAGCCTTGCAGGAAACCTTCACTGAGGCCCCAGGCTTCTTC 194
           | |||||| | ||| |||||||| |||||||| | ||||| |||  | |||||||||  |
Cp     453 CCAGTGAATGAGTAACCAGAAGTCTTGCAGGATATCTTCATTGAACCTCCAGGCTTCACC 394

Db     195 ACCTCAG 201
           |||||||
Cp     393 ACCTCAG 387


RESULT   12
LOCUS       H25625        488 bp    mRNA            EST       10-JUL-1995
DEFINITION  yl48g05.r1 Homo sapiens cDNA clone 161528 5' similar to gb:X72467
            IG KAPPA CHAIN PRECURSOR V-II REGION (HUMAN);.
ACCESSION   H25625
NID         g894748
KEYWORDS    EST.
SOURCE      human clone=161528 library=Soares breast 3NbHBst vector=pT7T3D
            (Pharmacia) with a modified polylinker host=DH10B (ampicillin
            resistant) primer=M13RP1 Rsite1=Not I Rsite2=Eco RI Adult human.
            1st strand cDNA was primed with a Not I -oligo(dT) primer [5'
            TGTTACCAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTT 3'],
            double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia),
            digested with Not I and cloned into the Not I and Eco RI sites of a
            modified pT7T3 vector (Pharmacia). Library went through one round
            of normalization to a Cot = 20. Library constructed by Bento Soares
            and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 488)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
```

```
COMMENT
           Contact: Wilson RK
           WashU-Merck EST Project
           Washington University School of Medicine
           4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
           Tel: 314 286 1800
           Fax: 314 286 1810
           Email: est@watson.wustl.edu
           High quality sequence stops: 362
           Source: IMAGE Consortium, LLNL
           This clone is available royalty-free through LLNL ; contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Homo sapiens"
                     /clone="161528"
BASE COUNT       99 a    130 c    128 g    127 t      4 others
ORIGIN

  Query Match           11.4%;  Score 81;  DB 16;  Length 488;
  Best Local Similarity  66.7%;  Pred. No. 4.51e-118;
  Matches   154;  Conservative    1;  Mismatches  76;  Indels   0;  Gaps   0;

Db    168 AGTTATTTGGATTGGTACCTGCAGAAGCCAGGACAGTCTCCACAGCTCCTGATCTATTTG 227
          |||||  || | |||||  | ||||||||||||||   || ||  | |    ||| |||
Qy     92 AGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAACCTTGGATTTATACC 151

Db    228 GCTTCTAAGCGGGCCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGAT 287
          | || || | :|| || || |||||||   | |||||||| |||| || || ||   |
Qy    152 ACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCT 211

Db    288 TTTACACTGAGAATCAGTAGAGTGGAGGCTGAGGATGTTGGCTTTTATTACTGCATGCAA 347
          |   | || | | |||| |||||||||||||| |||| || |  ||||||||||  |||
Qy    212 TACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAG 271

Db    348 ACTCTACAGACCCCGCTCACTTTCGGCGGAGGGACCAAGGTNGAGATNCAA 398
              |  | ||| | ||| ||||| || ||| ||||| | || ||  ||
Qy    272 TGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAAATAAAA 322


RESULT   13
LOCUS        AA300732      333 bp    mRNA            EST       18-APR-1997
DEFINITION   EST13847 Testis tumor Homo sapiens cDNA 5' end similar to similar
             to immunoglobulin kappa light chain (GB:S49006), mRNA sequence.
ACCESSION    AA300732
NID          g1953300
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
             Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
             Homo.
REFERENCE    1  (bases 1 to 333)
  AUTHORS    Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
             Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
             White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
             Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
             Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
             Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
             Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
             Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
             Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
             Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
             Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
             Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
             He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
             Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
             Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
             Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
             Fraser,C.M. and Venter,J.C.
  TITLE      Initial assessment of human gene diversity and expression patterns
             based upon 83 million nucleotides of cDNA sequence
  JOURNAL    Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE    96026280
COMMENT      Other_ESTs: THC87411
             Contact: Kerlavage, AR
             Bioinformatics
             The Institute for Genomic Research
             9712 Medical Center Drive, Rockville, MD 20850 USA
             Tel: 3018699056
             Fax: 3018699423
             Email: arkerlav@tigr.org
             For clone availability, additional sequence and expression
             information related to this EST, please check the TIGR Human Gene
             Index (http://www.tigr.org/tdb/hgi/hgi.html)
             Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .333
                     /organism="Homo sapiens"
                     /note="Organ: testis; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
                     /db_xref="ATCC (inhost):192041"
                     /db_xref="taxon:9606"
                     /clone_lib="Testis tumor"
                     /sex="male"
                     /dev_stage="adult".
     mRNA            <1. .>333
BASE COUNT       81 a     85 c     80 g     82 t      5 others
ORIGIN

  Query Match           11.2%;  Score 80;  DB 24;  Length 333;
  Best Local Similarity  67.7%;  Pred. No. 6.81e-116;
  Matches   149;  Conservative    1;  Mismatches  70;  Indels   0;  Gaps   0;

Db      3 CTGGTATCAGCAGAAACCAGGGAAAGCCCCTAAGCTCCTGATCTATGCTGCATCCAGTTT 62
          |||||||||||||||| |||||     |||| || |    ||| ||| |  ||||||   |
Qy    103 CTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCT 162

Db     63 GCAAAGTGGGGTCCCATCAAGGTTCAGCGGCAGTGGATCTGGGACAGATTTCACTCTCAC 122
          :     ||| |||||  |  | ||||| ||| |||| ||||||||   || | |||||||
Qy    163 KGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCAC 222

Db    123 CATCAGCAGCCTGCAGCCTGAAGATTTTGCAACTTACTATTGTCAACAGGCTAACAGTTT 182
          |||||||  || || ||||||||  ||| ||||| || || || |||    |   ||
Qy    223 AGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAG 282

Db    183 CCCGTGGACGTTCGGCCAAGGGACCAAGGTGGAAATCAAA 222
          |||    ||||||||    ||| ||||| ||||||| |||
Qy    283 CCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAAATAAAA 322


RESULT   14
LOCUS        AA301347      413 bp    mRNA            EST       18-APR-1997
DEFINITION   EST14279 Testis tumor Homo sapiens cDNA 5' end similar to similar
             to immunoglobulin kappa light chain, V region (GB:L01279), mRNA
             sequence.
ACCESSION    AA301347
NID          g1953680
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
             Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
             Homo.
REFERENCE    1  (bases 1 to 413)
  AUTHORS    Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
             Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
             White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
             Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
             Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
             Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
             Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
             Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
             Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
             Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
             Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
```

```
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES              Location/Qualifiers
    source            1. .413
                      /organism="Homo sapiens"
                      /note="Organ: testis; Vector: pBluescript SK-; Site_1:
                      EcoRI; Site_2: XhoI"
                      /db_xref="ATCC (inhost):191658"
                      /db_xref="taxon:9606"
                      /clone_lib="Testis tumor"
                      /sex="male"
                      /dev_stage="adult"
    mRNA              <1. .>413
BASE COUNT      100 a    110 c    105 g     98 t
ORIGIN

  Query Match           11.2%;  Score 80;  DB 24;  Length 413;
  Best Local Similarity  65.9%;  Pred. No. 6.81e-116;
  Matches   213;  Conservative    1;  Mismatches 104;  Indels   5;  Gaps   4;

Db      78 GACATTCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTGGGAGACAGAGT-CAC 136
           ||| | ||| ||||||||||||||| |   ||||||||||||||   || || |  || |||
Qy       4 GACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACAC 63

Db     137 CATCACTTGCCGGGCTGGGCAGCACATTAACACCTATTTAAATTGGTATCAACAGAAACC 196
            || |||||| ||||    |  ||  |  | |  ||  |  | ||||||| ||||| ||
Qy      64 AATGACTTGCAGGGCCACCC--CAAGTGTA-AGTTACATGCACTGGTATCAGCAGAAGCC 120

Db     197 AGGGAAAGCCCCTAAGCTCCTGATGTATGGTGCATCCAGTTTGCAAGGTGGGGTCCCTTC 256
           |||      |||| || |    ||| |||    ||||||   |:      ||| ||||||
Qy     121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180

Db     257 AAGGTTCAGTGGCGGTGGATCTGGGACAGATTTCACTCTCACCATCACCAGTCTGCAACC 316
            | ||||||||||||| |||||||   || | |||||||  ||| |||  || |  |
Qy     181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db     317 TGAAGATTTTGCAACTTACTACTTTCAGCAGAGTTACAGTCCTCATCAGAAGGTTCGGCC 376
           |||||||  ||| ||||| ||||  |||||| |     ||   |  |  | | |||||
Qy     241 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACG-TTCGGAG 299

Db     377 AAGGGACCAAGGTGGAAATCAAA 399
           ||| |||||  ||||||| |||
Qy     300 GGGGGTCCAAGCTGGAAATAAAA 322


RESULT   15
LOCUS       AI116236      116 bp    mRNA            EST       02-SEP-1998
DEFINITION  uc18e08.r1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone
            1398374 5' similar to gb:M87789 IG GAMMA-1 CHAIN C REGION (HUMAN);
            gb:X67210 M.musculus rearranged immunoglobulin gamma 2b heavy
            (MOUSE);, mRNA sequence.
ACCESSION   AI116236
NID         g3516560
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 116)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:910090
            Trace considered overall poor quality
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .116
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Organ: mammary gland; Vector: pT7T3D-Pac
                      (Pharmacia) with a modified polylinker; Site_1: Not I;
                      Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                      oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTTT
                      T 3']; double-stranded cDNA was ligated to Eco RI
                      adaptors (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of the modified pT7T3 vector.
                      RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                      constructed and normalized by Bento Soares and M.Fatima
                      Bonaldo."
                      /db_xref="taxon:10090"
                      /clone="1398374"
                      /clone_lib="Soares mouse mammary gland NbMMG"
                      /sex="male"
                      /tissue_type="mammary gland"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
BASE COUNT       38 a     32 c     25 g     21 t
ORIGIN

  Query Match           11.1%;  Score 79;  DB 17;  Length 116;
  Best Local Similarity  86.2%;  Pred. No. 1.02e-113;
  Matches    94;  Conservative    0;  Mismatches  15;  Indels   0;  Gaps   0;

Db       2 AGAAGTTCAAGAACAAGGCCACACTGACTGCAGACAAATCCTCCAGCACAGCCTACATGC 61
           ||||||||||  |||||||||| | |||| |||||| || ||||||||||||||||||
Qy     546 AGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATGG 605

Db      62 AACTCAGCAGCCTGACATCTGAGGACACTGCGGTCTATTACTGTGGAAG 110
           | |||  ||| |||||||||||||| |||| ||||||||||||| |||
Qy     606 AGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654


Search completed: Sat Jan  9 11:48:18 1999
Job time : 965 secs.
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Jan 9 13:21:31 1999; MasPar time 494.95 Seconds
1429.279 Million cell updates/sec
Tabular output not generated.

Title: >US-08-704-178-1
Description: (1-321) from US08704178.seq (2 of 4)
Perfect Score: 321
N.A. Sequence: 1 ATGGACCTGCAGCTGACCCA.....GGGTCCAAGCTGGAAATAAA 321
Comp: TACCTGGACGTCGACTGGGT.....CCCAGGTTCGACCTTTATTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 567134 seqs, 1101898692 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl55
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_om 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_vi
Database: genbank107
15:gb_ba1 16:gb_ba2 17:gb_htg 18:gb_in 19:gb_om 20:gb_ov
21:gb_pat 22:gb_ph 23:gb_pl1 24:gb_pl2 25:gb_pr1
26:gb_pr2 27:gb_pr3 28:gb_ro 29:gb_st 30:gb_sts 31:gb_sy
32:gb_un 33:gb_vi

Statistics: Mean 9.835; Variance 4.583; scale 2.146

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	314	97.8	711	21	I32406	Sequence 1 from patent	6.19e-249
2	314	97.8	711	31	S39590	anti-erbB2 imnmunotoxi	6.19e-249
3	270	84.1	315	21	I51657	Sequence 3 from patent	3.01e-209
4	270	84.1	315	21	A36642	Sequence 2 from Patent	3.01e-209
5	270	84.1	321	28	MMIG013	Mouse hybridoma 2d3 mR	3.01e-209
6	267	83.2	312	28	S48339	Ig V kappa =anti-idiot	1.51e-206
7	267	83.2	318	28	AF056217	Mus musculus monoclona	1.51e-206
8	266	82.9	1370	28	MUSIGKAAR	Mouse Ig kappa active	1.19e-205
9	265	82.6	345	28	MUSIGKVA2A	Mouse hybridoma Ig rea	9.46e-205
10	262	81.6	403	28	MUSIGKCNA	Mouse Ig rearranged ka	4.70e-202
11	262	81.6	403	21	I09200	Sequence 40 from Paten	4.70e-202
12	262	81.6	477	28	S76823	Ig V kappa =anti-sigma	4.70e-202
13	261	81.3	321	21	I27490	Sequence 4 from patent	3.72e-201
14	260	81.0	403	21	I05923	Sequence 39 from Paten	2.95e-200
15	258	80.4	309	28	MMIGVKP4	Mouse Ig active L-chai	1.84e-198
16	257	80.1	302	28	MUSIGKAFE	Mus musculus Mouse Ig	1.46e-197
17	256	79.8	403	21	I08812	Sequence 14 from Paten	1.15e-196
18	255	79.4	327	28	AF087025	Mus musculus hybrid 2-	9.11e-196
19	255	79.4	617	31	XXU05273	Synthetic construct cl	9.11e-196
20	254	79.1	384	28	MMU05285	Mus musculus BALB/c/J	7.20e-195
21	254	79.1	732	21	A51863	Sequence 27 from Paten	7.20e-195
22	253	78.8	321	28	MMU60466	Mus musculus Ig anti-D	5.69e-194
23	252	78.5	300	28	MUSIGKAFI	Mus musculus Ig active	4.49e-193
24	252	78.5	435	21	A17967	Variable region of the	4.49e-193
25	251	78.2	292	21	A07652	DNA sequence for monoc	3.54e-192
26	251	78.2	386	21	A20585	A5B7 humanised mouse m	3.54e-192
27	250	77.9	308	28	MMU55604	Mus musculus anti-DNA	2.80e-191
28	250	77.9	317	28	MMIGLT641	Mouse immunoglobulin v	2.80e-191
29	248	77.3	321	28	MMU223531	Mus musculus gene for	1.74e-189
30	248	77.3	321	28	MUSIGLBAA	Mus musculus Ig rearra	1.74e-189
31	248	77.3	349	28	MMIGG1LVR	M.musculus mRNA for Ig	1.74e-189
32	248	77.3	384	28	MUSIGKCF	Mouse Ig active kappa-	1.74e-189
33	248	77.3	384	28	MUSIGKCE	Mouse Ig active kappa-	1.74e-189
34	248	77.3	387	28	MUSIGKCH	Mouse Ig active kappa-	1.74e-189
35	247	76.9	318	28	MUSIKCVRV	Mouse mRNA for immunog	1.37e-188
36	246	76.6	402	28	S61341	Ig kappa =Ig kappa li	1.08e-187
37	245	76.3	321	28	MUSIKCVRR	Mouse mRNA for immunog	8.51e-187
38	245	76.3	617	28	MUSABCRB	Mus musculus antigen,	8.51e-187
39	244	76.0	309	28	MMIGVKPD	Mouse Ig active L-chai	6.70e-186
40	244	76.0	321	28	AF007246	Mus musculus domesticu	6.70e-186
41	244	76.0	639	21	I07388	Sequence 3 from Patent	6.70e-186
42	242	75.4	318	28	MUSIGMRNAD	Mus musculus immunoglo	4.15e-184
43	242	75.4	319	28	AF021856	Mus musculus clone B10	4.15e-184
44	242	75.4	321	28	MMU55427	Mus musculus Ig 3A6.A5	4.15e-184
45	242	75.4	639	28	S78551	Ig L=anti-alpha-fetopr	4.15e-184

ALIGNMENTS

RESULT 1
LOCUS I32406 711 bp DNA PAT 07-JAN-1997
DEFINITION Sequence 1 from patent US 5587458.
ACCESSION I32406
NID g1823197
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 711)
AUTHORS King,C.Richter, Kasprzyk,P.G. and Bird,R.E.
TITLE Anti-erbB-2 antibodies, combinations thereof, and therapeutic and diagnostic uses thereof
JOURNAL Patent: US 5587458-A 1 24-DEC-1996;
FEATURES Location/Qualifiers
source 1. .711
/organism="unknown"
BASE COUNT 175 a 182 c 190 g 164 t
ORIGIN

Query Match 97.8%; Score 314; DB 21; Length 711;
Best Local Similarity 99.4%; Pred. No. 6.19e-249;
Matches 319; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

```
Db      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT- 59
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db     60 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 120

Db    120 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTGGCTTCTGGAGTCCCTGC 179
          ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Qy    121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180

Db    180 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 239
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db    240 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 300

Db    300 GGGGTCCAAGCTGGAAATAAA 320
          |||||||||||||||||||||
Qy    301 GGGGTCCAAGCTGGAAATAAA 321


RESULT    2
LOCUS       S39590        711 bp                    SYN       10-FEB-1993
DEFINITION  anti-erbB2 imnmunotoxin antigen binding region [mice, Other
            Synthetic Recombinant Partial, 711 nt].
ACCESSION   S39590
NID         g251113
KEYWORDS    .
SOURCE      Mus sp.
  ORGANISM  Mus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 711)
  AUTHORS   Batra,J.K., Kasprzyk,P.G., Bird,R.E., Pastan,I. and King,C.R.
  TITLE     Recombinant anti-erbB2 immunotoxins containing Pseudomonas exotoxin
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 89 (13), 5867-5871 (1992)
  MEDLINE   92335198
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 108547] from the original journal article.
            This sequence comes from Fig. 1.
FEATURES             Location/Qualifiers
     source          1. .711
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
     gene            1. .711
                     /partial
                     /gene="anti-erbB2 imnmunotoxin antigen binding region"
     CDS             1. .711
                     /partial
                     /gene="anti-erbB2 imnmunotoxin antigen binding region"
                     /note="This sequence comes from Fig. 1."
                     /codon_start=1
                     /db_xref="PID:g251114"
                     /translation="MDLQLTQSPAILSASPGEKVTMTCRATPSVSYMHWYQQKPGSSP
                     KPWIYTTSNLASGVPARFSGGGSGTSYSLTVSRVEAEDAATYYCQQWSRSPPTFGGGS
                     KLEIKGSTSGSGKSSEGKGVQLQESGPEVVKPGGSMKISCKTSGYSFTGHTMNWVKQS
                     HGKNLEWIGLINPYNGDTNYNQKFKGKATFTVDKSSSTAYMELLSLTSEDSAVYYCAR
                     RVTDWYFDVWGAGTTVTVS"
BASE COUNT      175 a    182 c    190 g    164 t
ORIGIN

  Query Match           97.8%;  Score 314;  DB 31;  Length 711;
  Best Local Similarity  99.4%;  Pred. No. 6.19e-249;
  Matches   319;  Conservative    1;  Mismatches   0;  Indels   1;  Gaps   1;

Db      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT- 59
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db     60 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 120

Db    120 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTGGCTTCTGGAGTCCCTGC 179
          |||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Qy    121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180

Db    180 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db    240 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 300

Db    300 GGGGTCCAAGCTGGAAATAAA 320
          |||||||||||||||||||||
Qy    301 GGGGTCCAAGCTGGAAATAAA 321


RESULT    3
LOCUS       I51657        315 bp    DNA             PAT       12-AUG-1997
DEFINITION  Sequence 3 from patent US 5645817.
ACCESSION   I51657
NID         g2472858
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 315)
  AUTHORS   Seemann,G. and Bosslet,K.
  TITLE     Granulocyte-binding antibody constructs, their preparation and use
  JOURNAL   Patent: US 5645817-A 3 08-JUL-1997;
FEATURES             Location/Qualifiers
     source          1. .315
                     /organism="unknown"
BASE COUNT       72 a     92 c     80 g     71 t
ORIGIN

  Query Match           84.1%;  Score 270;  DB 21;  Length 315;
  Best Local Similarity  93.3%;  Pred. No. 3.01e-209;
  Matches   294;  Conservative    1;  Mismatches  19;  Indels   1;  Gaps   1;

Db      1 GACATTCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CAC 59
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||| |||
Qy      4 GACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACAC 63

Db     60 AATGACTTGCAGGGCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGG 119
          |||||||||||||||| | ||||||||||||||||||||||| ||||||||||||||
Qy     64 AATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGG 123

Db    120 ATCCTCCCCCAAACCCTGGATTTATGCCACATCCAACCTGGCTTCTGGAGTCCCTGCTCG 179
          |||||||||||||| ||||||||| |||||||||||||:||||||||||||||||||||
Qy    124 ATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCG 183

Db    180 CTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCATCAGAGTGGAGGCTGA 239
          |||||||||| |||||||||||||||||||||||||||| ||| ||||||||||||||
Qy    184 CTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGA 243

Db    240 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTAGTAACCCGCTCACGTTCGGTGCTGG 299
          |||||||||||||||||||||||||||||||||||| ||| ||| | |||||||| |  ||
Qy    244 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGG 303

Db    300 GACCAAGCTGGAGAT 314
          | |||||||||| ||
Qy    304 GTCCAAGCTGGAAAT 318


RESULT    4
LOCUS       A36642        315 bp    DNA             PAT       05-MAR-1997
DEFINITION  Sequence 2 from Patent EP0585570.
ACCESSION   A36642
NID         g2293945
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 315)
  AUTHORS   Seemann,G. and Bosslet,K.
  TITLE     Antigranulocyte antibody construct, preparation and use thereof
  JOURNAL   Patent: EP 0585570-A 2 09-MAR-1994;
            BEHRINGWERKE AG (DE)
```

```
COMMENT     Other publication JP 6205693 940726
            Other publication CA 2101868 940206
            Other publication AU 4441393 940210
            Other publication AU 666029 960125
            Other publication DE 4225853 940210.
FEATURES             Location/Qualifiers
     source          1. .315
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /cell_type="hybridoma/myeloma"
BASE COUNT       72 a     92 c     80 g     71 t
ORIGIN

  Query Match           84.1%;  Score 270;  DB 21;  Length 315;
  Best Local Similarity  93.3%;  Pred. No. 3.01e-209;
  Matches   294;  Conservative    1;  Mismatches  19;  Indels   1;  Gaps   1;

Db       1 GACATTCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CAC 59
           ||| | |||||||||||||||||||||||||||||||||||||||||||||||||| |||
Qy       4 GACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACAC 63

Db      60 AATGACTTGCAGGGCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGG 119
           |||||||||||||||| | |||||||||||||||||||||||| ||||||||||||||
Qy      64 AATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGG 123

Db     120 ATCCTCCCCCAAACCCTGGATTTATGCCACATCCAACCTGGCTTCTGGAGTCCCTGCTCG 179
           |||||||||||||||| ||||||||| |||||||||||||:||||||||||||||||||
Qy     124 ATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCG 183

Db     180 CTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCATCAGAGTGGAGGCTGA 239
           |||||||||| |||||||||||||||||||||||||| ||| ||||||||||||||||
Qy     184 CTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGA 243

Db     240 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTAGTAACCCGCTCACGTTCGGTGCTGG 299
           |||||||||||||||||||||||||||||||||||| ||| ||| | ||||||||| |  ||
Qy     244 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGG 303

Db     300 GACCAAGCTGGAGAT 314
           | |||||||||| ||
Qy     304 GTCCAAGCTGGAAAT 318


RESULT    5
LOCUS       MMIG013       321 bp    RNA             ROD       06-AUG-1995
DEFINITION  Mouse hybridoma 2d3 mRNA for immunoglobuin kappa light chain V
            region.
ACCESSION   X58586 Y00794
NID         g51562
KEYWORDS    Ig kappa light chain; Ig variable region; immunoglobulin.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 321)
  AUTHORS   Meek,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-1989) K. Meek, University of Texas Southwestern
            Medical School, 5323 Harry Hines Blud. Microbiology, Dallas Texas
            75235, USA
REFERENCE   2  (bases 1 to 321)
  AUTHORS   Meek,K., Hasemann,C., Pollok,B., Alkan,S.S., Brait,M., Slaoui,M.,
            Urbain,J. and Capra,J.D.
  TITLE     Structural characterization of antiidiotypic antibodies. Evidence
            that Ab2s are derived from the germline differently than Ab1s
  JOURNAL   J. Exp. Med. 169 (2), 519-533 (1989)
  MEDLINE   89094248
FEATURES             Location/Qualifiers
     source          1. .321
                     /organism="Mus musculus"
                     /strain="Balb/c"
                     /db_xref="taxon:10090"
                     /cell_line="Hybridoma 2D3-K"
     V_region        <1. .>321
                     /gene="IG kappa light chain"
                     /note="variable region"
                     /product="IG kappa light chain"
     mRNA            <1. .>321
                     /gene="IG kappa light chain"
                     /note="variable region"
                     /evidence=experimental
     gene            1. .321
                     /gene="IG kappa light chain"
     CDS             <1. .>321
                     /gene="IG kappa light chain"
                     /note="variable region"
                     /codon_start=1
                     /product="IG kappa light chain"
                     /db_xref="PID:g938245"
                     /translation="QIVLSQSPAILSASPGEKVTMTCRASSSVSYMHWYQQKPGSSPK
                     PWISATSNLASGVPARFSGSGSGTSYSLTISRVEAEDAATYYCHQWSSNPPTFGGGTK
                     LEIKR"
BASE COUNT       78 a     92 c     80 g     71 t
ORIGIN

  Query Match           84.1%;  Score 270;  DB 28;  Length 321;
  Best Local Similarity  94.8%;  Pred. No. 3.01e-209;
  Matches   289;  Conservative    1;  Mismatches  14;  Indels   1;  Gaps   1;

Db      14 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CACAATGACTTGCAGG 72
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Qy      17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db      73 GCCAGCTCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCAAAA 132
           |||| | |||||||||||||||||||||||||||||||||||||||||||||||||| |||
Qy      77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db     133 CCCTGGATTTCTGCCACATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGT 192
           || ||||||| | |||||||||||||:||||||||||||||||||||||||||||||| ||
Qy     137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db     193 GGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 252
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy     197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db     253 TATTACTGCCACCAGTGGAGTAGTAACCCCCCAACGTTCGGAGGGGGGACCAAGCTGGAA 312
           |||||||||||| ||||||||| ||| ||| || ||||||||||||||| ||||||||||
Qy     257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAA 316

Db     313 ATAAA 317
           |||||
Qy     317 ATAAA 321


RESULT    6
LOCUS       S48339        312 bp    mRNA            ROD       14-JAN-1993
DEFINITION  Ig V kappa =anti-idiotypic Fab [mice, mRNA Partial, 312 nt].
ACCESSION   S48339
NID         g257749
KEYWORDS    .
SOURCE      Mus sp.
  ORGANISM  Mus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 312)
  AUTHORS   Kasai,Y., Herlyn,D., Sperlagh,M., Maruyama,H., Matsushita,S. and
            Linnenbach,A.J.
  TITLE     Molecular cloning of murine monoclonal anti-idiotypic Fab
  JOURNAL   J. Immunol. Methods 155 (1), 77-89 (1992)
  MEDLINE   93017981
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 118382] from the original journal article.
            This sequence comes from Fig. 5.
FEATURES             Location/Qualifiers
     source          1. .312
```

```
                  /organism="Mus sp."
                  /db_xref="taxon:10095"
     gene         1. .312
                  /partial
                  /note="anti-idiotypic Fab"
                  /gene="Ig V<down>&kgr;</down>"
     CDS          1. .312
                  /partial
                  /gene="Ig V<kappa>"
                  /note="anti-idiotypic Fab; This sequence comes from Fig.
                  5"
                  /codon_start=1
                  /db_xref="PID:g257750"
                  /translation="VMTQSPAILSASPGEKVTMTCRASSSVSYMHWYQQKPGSSPKPW
                  IYATSNLASGVPARFSGSGSGTSYSLTISRVEAEDAATYYCQQWSSNPWTFGGGTKLE
                  IK"
BASE COUNT       76 a     88 c     80 g     68 t
ORIGIN

  Query Match           83.2%;  Score 267;  DB 28;  Length 312;
  Best Local Similarity  93.8%;  Pred. No. 1.51e-206;
  Matches   289;  Conservative    1;  Mismatches  17;  Indels   1;  Gaps   1;

Db       5 TGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CACAATGACTTGC 63
           |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Qy      14 TGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGC 73

Db      64 AGGGCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGGATCCTCCCCC 123
           ||||||| | ||||||||||||||||||||||| ||||||||||||||||||||||||||
Qy      74 AGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCC 133

Db     124 AAACCCTGGATTTATGCCACATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGC 183
           ||||| |||||||||| |||||||||||||:||||||||||||||||||||||||||||||
Qy     134 AAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGC 193

Db     184 AGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCC 243
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Qy     194 GGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCC 253

Db     244 ACTTATTACTGCCAGCAGTGGAGTAGTAACCCATGGACGTTCGGTGGAGGCACCAAGCTG 303
           |||||||||||||||||||||||| ||| ||||   |||||||| || ||  ||||||||
Qy     254 ACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTG 313

Db     304 GAAATCAA 311
           ||||| ||
Qy     314 GAAATAAA 321


RESULT    7
LOCUS       AF056217      318 bp    mRNA            ROD       13-APR-1998
DEFINITION  Mus musculus monoclonal antibody aH7:38 IgG1 light chain mRNA,
            partial cds.
ACCESSION   AF056217
NID         g3044131
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 318)
  AUTHORS   Sandstrom,P., Johansson,A., Ullen,A., Behravan,G. and Stigbrand,T.
  TITLE     Differences in sequence and affinity between three monoclonal
            anti-idiotypic antibodies against one anti-placental alkaline
            phosphatase idiotypic antibody
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 318)
  AUTHORS   Sandstrom,P., Johansson,A., Ullen,A., Behravan,G. and Stigbrand,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-1998) Immunology, Umea University, Umea 90185,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .318
                     /organism="Mus musculus"
                     /strain="Balb/c"
                     /db_xref="taxon:10090"
                     /cell_line="hybridoma"
     CDS             <1. .>318
                     /note="anti-idiotypic antibody against anti-placental
                     alkaline phosphatase antibody; variable region"
                     /codon_start=1
                     /product="monoclonal antibody aH7:38 IgG1 light chain"
                     /db_xref="PID:g3044132"
                     /translation="DIQLTQSPAILSASPGEKVSMTCRASSSVSYIHWYQQKPGSSPK
                     PWIYATSNLASGVPVRFSGSGSGTSYSLTISRVEAEDAATYYCQQWSGNPLTFGAGTK
                     LEIK"
BASE COUNT       74 a     91 c     81 g     72 t
ORIGIN

  Query Match           83.2%;  Score 267;  DB 28;  Length 318;
  Best Local Similarity  92.5%;  Pred. No. 1.51e-206;
  Matches   294;  Conservative    1;  Mismatches  22;  Indels   1;  Gaps   1;

Db       1 GACATTCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CTC 59
           ||| | |||||||||||||||||||||||||||||||||||||||||||||||||| | |
Qy       4 GACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACAC 63

Db      60 AATGACTTGCAGGGCCAGCTCAAGTGTAAGTTACATACACTGGTACCAGCAGAAGCCAGG 119
           |||||||||||||||| | |||||||||||||||| |||||||| ||||||||||||||
Qy      64 AATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGG 123

Db     120 ATCCTCCCCCAAACCCTGGATTTATGCCACATCCAACCTGGCTTCTGGAGTCCCTGTTCG 179
           |||||||||||||||| |||||||| |||||||||||||:|||||||||||||||| |||
Qy     124 ATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCG 183

Db     180 CTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGA 239
           |||||||||| |||||||||||||||||||||||||||||||||| ||||||||||||||
Qy     184 CTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGA 243

Db     240 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTGGTAACCCGCTCACGTTCGGTGCTGG 299
           |||||||||||||||||||||||||||||||| ||| ||| | ||||||||| |  ||
Qy     244 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGG 303

Db     300 GACCAAGCTGGAGATCAA 317
           | |||||||||| || ||
Qy     304 GTCCAAGCTGGAAATAAA 321


RESULT    8
LOCUS       MUSIGKAAR    1370 bp    DNA             ROD       05-FEB-1992
DEFINITION  Mouse Ig kappa active V-region from 70Z/3 cells.
ACCESSION   K01641
NID         g196460
KEYWORDS    C-region; V-region; immunoglobulin light chain;
            immunoglobulin-kappa.
SOURCE      Mouse 70Z/3 cell DNA.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 1370)
  AUTHORS   Parslow,T.G., Blair,D.L., Murphy,W.J. and Granner,D.K.
  TITLE     Structure of the 5' ends of immunoglobulin genes: A novel conserved
            sequence
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 81, 2650-2654 (1984)
  MEDLINE   84194064
COMMENT     Corrections to the sequence were sent by Dr. Richard E. Manroe and
            Dr. Tristram G. Parslow on Dec. 9, 1991.
FEATURES             Location/Qualifiers
     source          1. .1370
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     exon            <776. .824
                     /note="Ig kappa chain variable region"
                     /number=1
     CDS             join(776. .824,1003. .1337)
                     /partial
```

```
                  /note="Ig kappa chain variable region"
                  /codon_start=1
                  /db_xref="PID:g196461"
                  /translation="MDFQVQIFSFLLISASVIMSRGQIVLSQSPAILSASPGEKVTMT
                  CRASSSVSYMHWYQQKPGSSPKPWIYATSNLASGVPARFSGSGSGTSYSLTISRVEAE
                  DAATYYCQQWSSNPRTFGGGTKLEIK"
     intron       825. .1002
                  /note="kappa cds intron"
     exon         1003. .>1337
                  /note="Ig kappa chain variable region"
                  /number=2
     intron       1338. .>1369
                  /note="kappa cds intron J-C"
BASE COUNT      404 a    273 c    253 g    440 t
ORIGIN

  Query Match           82.9%;  Score 266;  DB 28;  Length 1370;
  Best Local Similarity  94.1%;  Pred. No. 1.19e-205;
  Matches   287;  Conservative    1;  Mismatches  16;  Indels   1;  Gaps   1;

Db   1033 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CACAATGACTTGCAGG 1091
         ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db   1092 GCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGGATCCTCCCCCAAA 1151
         |||| | ||||||||||||||||||||||| ||||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db   1152 CCATGGATTTATGCCACATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGT 1211
         || |||||||| ||||||||||||:|||||||||||||||||||||||||||||||||| ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db   1212 GGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 1271
         |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db   1272 TATTACTGCCAGCAGTGGAGTAGTAACCCACGGACGTTCGGTGGAGGCACCAAGCTGGAA 1331
         |||||||||||||||||||| ||| |||||  |||||||| || ||  ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAA 316

Db   1332 ATCAA 1336
         || ||
Qy    317 ATAAA 321


RESULT    9
LOCUS       MUSIGKVA2A     345 bp    mRNA            ROD       23-MAR-1994
DEFINITION  Mouse hybridoma Ig rearranged kappa-chain mRNA V-region, partial
            cds.
ACCESSION   M97864
NID         g197422
KEYWORDS    immunoglobulin light chain; immunoglobulin-kappa; processed gene;
            variable region.
SOURCE      Mus musculus (strain BALB/c, sub_species domesticus) NS1-BALB/c
            fusion hybridoma cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 345)
  AUTHORS   Lohman,K.L., Kieber-Emmons,T. and Kennedy,R.C.
  TITLE     Molecular characterization and structural modeling of
            immunoglobulin variable regions from murine monoclonal antibodies
            specific for hepatitis B virus surface antigen
  JOURNAL   Mol. Immunol. 30 (14), 1295-1306 (1993)
  MEDLINE   94019421
FEATURES             Location/Qualifiers
     source          1. .345
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
BASE COUNT       78 a    101 c     89 g     77 t
ORIGIN
```

```
  Query Match           82.6%;  Score 265;  DB 28;  Length 345;
  Best Local Similarity  93.5%;  Pred. No. 9.46e-205;
  Matches   288;  Conservative    1;  Mismatches  18;  Indels   1;  Gaps   1;

Db      7 CTCACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CACAATGACTTG 65
          || ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Qy     13 CTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTG 72

Db     66 CAGGGCCAGTTCAAGTGTAGGTTCCATACACTGGTACCAGCAGAAGCCAGGATCCTCCCC 125
          ||||||||   |||||||| ||| ||| |||||||| ||||||||||||||||||||||
Qy     73 CAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCC 132

Db    126 CAAACCCTGGATTTATGCCACATCCAACCTGGCGTCTGGAGTCCCTGCTCGCTTCAGTGG 185
          |||||| ||||||||| |||||||||||||:|| ||||||||||||||||||||||||||
Qy    133 CAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGG 192

Db    186 CAGTGGGTCCGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGC 245
          | ||||||| |||||||||||||||||||| ||||||||||||||||||||||||||||
Qy    193 CGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGC 252

Db    246 CACTTTTTACTGCCAGCAGTGGAGTAGTAACCCACCCACGTTCGGAGGGGGGTCCAGGCT 305
          ||||| |||||||||||||||||| ||| |||||||||||||||||||||||||| |||
Qy    253 CACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCT 312

Db    306 GGAAATAA 313
          ||||||||
Qy    313 GGAAATAA 320


RESULT   10
LOCUS       MUSIGKCNA     403 bp    mRNA            ROD       26-MAR-1994
DEFINITION  Mouse Ig rearranged kappa-chain mRNA VJ5.
ACCESSION   M17954
NID         g197015
KEYWORDS    C-region; V-region; immunoglobulin light chain;
            immunoglobulin-kappa; processed gene.
SOURCE      Mus musculus cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 403)
  AUTHORS   Liu,A.Y., Robinson,R.R., Murray,E.D.Jr. ., Ledbetter,J.A.,
            Hellstroem,I. and Hellstroem,K.E.
  TITLE     Production of a mouse-human chimeric monoclonal antibody to CD20
            with potent Fc-dependent biologic activity
  JOURNAL   J. Immunol. 139, 3521-3526 (1987)
  MEDLINE   88060445
COMMENT     Draft entry and computer-readable sequence [1] kindly submitted by
            A.Y.Liu, 02-FEB-1987.
FEATURES             Location/Qualifiers
     source          1. .403
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /cell_line="2H7"
     CDS             20. .>403
                     /note="V-J-5 region"
                     /codon_start=1
                     /product="immunoglobulin kappa-chain"
                     /db_xref="PID:g466303"
                     /translation="MDFQVQIFSFLLISASVIIARGQIVLSQSPAILSASPGEKVTMT
                     CRASSSVSYMHWYQQKPGSSPKPWIYAPSNLASGVPARFSGSGSGTSYSLTISRVEAE
                     DAATYYCQQWSFNPPTFGAGTKLELK"
     misc_recomb     370. .371
                     /note="V-region end/J-5-region start"
                     /organism="Mus musculus"
BASE COUNT      100 a    112 c     93 g     98 t
ORIGIN

  Query Match           81.6%;  Score 262;  DB 28;  Length 403;
  Best Local Similarity  94.3%;  Pred. No. 4.70e-202;
  Matches   282;  Conservative    1;  Mismatches  15;  Indels   1;  Gaps   1;
```

```
Db      99 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CACAATGACTTGCAGG 157
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Qy      17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db     158 GCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGGATCCTCCCCCAAA 217
           |||| | |||||||||||||||||||||||| ||||||||||||||||||||||||||||
Qy      77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db     218 CCCTGGATTTATGCCCCATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGT 277
           || ||||||||| || ||||||||||:|||||||||||||||||||||||||||||||||
Qy     137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db     278 GGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 337
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Qy     197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db     338 TATTACTGCCAGCAGTGGAGTTTTAACCCACCCACGTTCGGTGCTGGGACCAAGCTGGA 396
           |||||||||||||||||||  || |||||||||||||| |  ||| ||||||||||
Qy     257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGA 315


RESULT   11
LOCUS       I09200        403 bp                    PAT       14-NOV-1994
DEFINITION  Sequence 40 from Patent WO 8900999.
ACCESSION   I09200
NID         g588127
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 403)
  AUTHORS   Robinson,R.R., Liu,A.Y., Horwitz,A.H., Wall,R. and Better,M.
  TITLE     MODULAR ASSEMBLY OF ANTIBODY GENES, ANTIBODIES PREPARED THEREBY AND
            USE
  JOURNAL   Patent: WO 8900999-A 40 09-FEB-1989;
FEATURES             Location/Qualifiers
     source          1. .403
                     /organism="unknown"
BASE COUNT      100 a    113 c     92 g     98 t
ORIGIN

  Query Match           81.6%;  Score 262;  DB 21;  Length 403;
  Best Local Similarity  94.3%;  Pred. No. 4.70e-202;
  Matches   282;  Conservative    1;  Mismatches  15;  Indels   1;  Gaps   1;

Db      99 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CACAATGACTTGCAGG 157
           |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Qy      17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db     158 GCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGGATCCTCCCCCAAA 217
           |||| | |||||||||||||||||||||||| ||||||||||||||||||||||||||||
Qy      77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db     218 CCCTGGATTTATGCCCCATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGT 277
           || ||||||||| || ||||||||||:|||||||||||||||||||||||||||||||||
Qy     137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db     278 GGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 337
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Qy     197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db     338 TATTACTGCCAGCAGTGGAGTTTTAACCCACCCACGTTCGGTGCTGGGACCAAGCTGGA 396
           |||||||||||||||||||  || |||||||||||||| |  ||| ||||||||||
Qy     257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGA 315


RESULT   12
LOCUS       S76823        477 bp    mRNA            ROD       26-JUL-1995
DEFINITION  Ig V kappa =anti-sigma receptor Ig light chain [mice, hybridoma
            cell line 10G9, mRNA Partial, 477 nt].
ACCESSION   S76823
NID         g913971
KEYWORDS    .
SOURCE      Mus sp. hybridoma cell line 10G9.
  ORGANISM  Mus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 477)
  AUTHORS   Kobayashi,T., Togashi,S., Itoh,N. and Kumanishi,T.
  TITLE     Molecular cloning of cDNAs for immunoglobulin variable regions of a
            monoclonal anti-idiotypic antibody specific for sigma receptors
  JOURNAL   J. Neuroimmunol. 57 (1-2), 129-135 (1995)
  MEDLINE   95221566
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 163632] from the original journal article.
            This sequence comes from Fig. 4.
FEATURES             Location/Qualifiers
     source          1. .477
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
     gene            52. .477
                     /partial
                     /note="anti-sigma receptor Ig light chain"
                     /gene="Ig V&kgr;"
     CDS             52. .477
                     /partial
                     /gene="Ig V<kappa>"
                     /note="monoclonal anti-idiotypic antibody specific for
                     sigma receptors; This sequence comes from Fig. 4"
                     /codon_start=1
                     /product="anti-sigma receptor Ig light chain"
                     /db_xref="PID:g913972"
                     /translation="MDFQVQIFSFLLISASVIISRGQIVLSQSPAILSASPGEKVTMT
                     CRASSSVSYIHWYQQKPGSSPKPWILATSNLASGVPARFSGSGSGTSYSLTISRMEAE
                     DAATYYCQQWSRNLYTFGGGTKLEIKRLMLHQLYPSSHHP"
BASE COUNT      125 a    128 c    106 g    118 t
ORIGIN

  Query Match           81.6%;  Score 262;  DB 28;  Length 477;
  Best Local Similarity  93.4%;  Pred. No. 4.70e-202;
  Matches   285;  Conservative    1;  Mismatches  18;  Indels   1;  Gaps   1;

Db     131 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CACAATGACTTGCAGG 189
           |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Qy      17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db     190 GCCAGCTCAAGTGTAAGTTACATTCACTGGTACCAGCAGAAGCCAGGATCCTCCCCCAAA 249
           |||| | |||||||||||||||| |||||||| ||||||||||||||||||||||||||||
Qy      77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db     250 CCCTGGATTTTAGCCACATCCAACCTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGT 309
           || |||||||   |||||||||||||:|||||||||||||||||||||||||||||||||
Qy     137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db     310 GGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAATGGAGGCTGAAGATGCTGCCACT 369
           |||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||||
Qy     197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db     370 TATTACTGCCAGCAGTGGAGTCGTAACCTGTACACGTTCGGCGGGGGGACCAAGCTGGAA 429
           |||||||||||||||||||||||||| ||    |||||||| |||||| ||||||||||
Qy     257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAA 316

Db     430 ATAAA 434
           |||||
Qy     317 ATAAA 321


RESULT   13
LOCUS       I27490        321 bp    DNA             PAT       16-OCT-1996
DEFINITION  Sequence 4 from patent US 5565332.
ACCESSION   I27490
NID         g1818266
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
```

```
          Unclassified.
REFERENCE   1  (bases 1 to 321)
  AUTHORS   Hoogenboom,H.R.J.M., Baier,M., Jespers,L.S.A.T. and Winter,G.P.
  TITLE     Production of chimeric antibodies - a combinatorial approach
  JOURNAL   Patent: US 5565332-A 4 15-OCT-1996;
FEATURES             Location/Qualifiers
     source          1. .321
                     /organism="unknown"
BASE COUNT       75 a      95 c      80 g      71 t
ORIGIN

  Query Match           81.3%;  Score 261;  DB 21;  Length 321;
  Best Local Similarity  91.5%;  Pred. No. 3.72e-201;
  Matches   291;  Conservative    1;  Mismatches  25;  Indels   1;  Gaps   1;

Db       1 GACATTGAGCTCACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGGGAAGGT-CAC 59
           ||| |  |||| ||||||||||||||||||||||||||||||||||||||||| |||||| |||
Qy       4 GACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACAC 63

Db      60 AATGACTTGTAGGGCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGG 119
           ||||||||| ||||||| | ||||||||||||||||||||||| ||||||||||||||
Qy      64 AATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGG 123

Db     120 ATCCTCCCCCAAACCCTGGATTTATGCCACATCCAACCTGGCTTCTGGAGTCCCTACTCG 179
           |||||||||||||| |||||||| |||||||||||||:|||||||||||||| ||||
Qy     124 ATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCG 183

Db     180 CTTCAGTGGCACTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGGGTGGAGGCTGA 239
           |||||||||  |||||||||||||||||||||||||||| ||||||| ||||||||||
Qy     184 CTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGA 243

Db     240 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAACCCATTCACGTTCGGCTCGGG 299
           |||||||||||||||||||||||||||||||||||| ||||  |||||||||   |||
Qy     244 AGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGG 303

Db     300 CACCAAGCTGGAAATCAA 317
           |||||||||||||  ||
Qy     304 GTCCAAGCTGGAAATAAA 321


RESULT   14
LOCUS       I05923       403 bp                    PAT       14-NOV-1994
DEFINITION  Sequence 39 from Patent EP 0274394.
ACCESSION   I05923
NID         g590877
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 403)
  AUTHORS   Robinson,R.R., Liu,A.Y., Hellstrom,K.E., Hellstrom,I. and
            Ledbetter,J.A.
  TITLE     Chimeric antibody with specificity to human B cell surface antigen
  JOURNAL   Patent: EP 0274394-A2 39 13-JUL-1988;
FEATURES             Location/Qualifiers
     source          1. .403
                     /organism="unknown"
BASE COUNT      100 a     111 c      93 g      99 t
ORIGIN

  Query Match           81.0%;  Score 260;  DB 21;  Length 403;
  Best Local Similarity  94.0%;  Pred. No. 2.95e-200;
  Matches   281;  Conservative    1;  Mismatches  16;  Indels   1;  Gaps   1;

Db      99 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGT-CACAATGACTTGCAGG 157
           ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Qy      17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db     158 GCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGGATCCTCCCCCAAA 217
           |||| | |||||||||||||||||||||||| ||||||||||||||||||||||||||||
Qy      77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db     218 CCCTGGATTTATGCCCCATCCAACCTGGCTTCTGGAGTCCCTGCTCGTTTCAGTGGCAGT 277
           || |||||||| || ||||||||||:|||||||||||||||||||| |||||||| ||
Qy     137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db     278 GGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 337
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy     197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db     338 TATTACTGCCAGCAGTGGAGTTTTAACCCACCCACGTTCGGTGCTGGGACCAAGCTGGA 396
           |||||||||||||||||||  || |||||||||||||| |  ||| ||||||||||
Qy     257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGA 315


RESULT   15
LOCUS       MMIGVKP4      309 bp    RNA             ROD       04-AUG-1992
DEFINITION  Mouse Ig active L-chain V-region: anti-phOx mRNA (NQ18.36.44L).
ACCESSION   X06517 Y00010
NID         g52547
KEYWORDS    Ig kappa light chain; Ig light chain; variable region.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 309)
  AUTHORS   Berek,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-DEC-1987) Berek C., Institut fuer Genetik, 5 Koeln
            41, Weyertal 121 W-Germany
REFERENCE   2  (bases 1 to 309)
  AUTHORS   Berek,C., Jarvis,J.M. and Milstein,C.
  TITLE     Activation of memory and virgin B cell clones in hyperimmune
            animals
  JOURNAL   Eur. J. Immunol. 17 (8), 1121-1129 (1987)
  MEDLINE   87304452
FEATURES             Location/Qualifiers
     source          1. .309
                     /organism="Mus musculus"
                     /strain="Balb/c"
                     /db_xref="taxon:10090"
                     /cell_type="hybridoma"
     misc_feature    1. .309
                     /note="transcribed region"
BASE COUNT       69 a      91 c      76 g      71 t       2 others
ORIGIN

  Query Match           80.4%;  Score 258;  DB 28;  Length 309;
  Best Local Similarity  93.6%;  Pred. No. 1.84e-198;
  Matches   278;  Conservative    1;  Mismatches  17;  Indels   1;  Gaps   1;

Db      14 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGCAGGT-CACAATGACTTGCAGG 72
           ||||||||||||||||||||||||||||||||||||||||||| |||| ||||||||||||||
Qy      17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db      73 GCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGGATCCTCCCCCAAA 132
           |||| | |||||||||||||||||||||||| ||||||||||||||||||||||||||||
Qy      77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db     133 CCCTGGATTTATGCCACATCCAACNTGGCTTCTGGAGTCCCNGCTCGCTTCAGTGGCAGT 192
           || |||||||| |||||||||| |:|||||||||||||| ||||||||||||||| ||
Qy     137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db     193 GGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 252
           ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Qy     197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db     253 TATTACTGCCAGCAGTGGAGTAGTAACCCACTCACGTTCGGTGCTGGGACCAAGCTG 309
           |||||||||||||||||||| ||| ||||| ||||||||| |  ||| ||||||||
Qy     257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTG 313


Search completed: Sat Jan  9 13:29:52 1999
Job time : 501 secs.
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Jan 9 14:31:15 1999; MasPar time 60.25 Seconds
724.916 Million cell updates/sec

Tabular output not generated.

Title: >US-08-704-178-1
Description: (1-321) from US08704178.seq (2 of 4)
Perfect Score: 321
N.A. Sequence: 1 ATGGACCTGCAGCTGACCCA.....GGGTCCAAGCTGGAAATAAA 321
Comp: TACCTGGACGTCGACTGGGT.....CCCAGGTTCGACCTTTATTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 8.055; Variance 4.678; scale 1.722

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	314	97.8	711	9	Q55180	Sequence encoding the	2.02e-204
2	314	97.8	711	29	T65006	Single-chain anti-erb	2.02e-204
3	312	97.2	711	18	T17728	Anti-erbB2 scFv cDNA.	5.83e-203
4	264	82.2	384	11	Q65630	Murine variable regio	5.34e-168
5	262	81.6	318	34	T79900	Anti-Factor IX MAb ch	1.52e-166
6	262	81.6	321	34	T77377	Mouse anti-human Fact	1.52e-166
7	262	81.6	335	34	T79899	Anti-Factor IX MAb BC	1.52e-166
8	262	81.6	403	2	N70972	2H7 VL sequence in wh	1.52e-166
9	262	81.6	403	1	N91147	2H7 Vh sequence.	1.52e-166
10	262	81.6	426	30	T51043	Coding sequence for l	1.52e-166
11	262	81.6	426	32	T36317	2H7 antibody light ch	1.52e-166
12	262	81.6	426	31	T70869	2H7 light chain varia	1.52e-166
13	262	81.6	426	40	V03927	Mouse 2H7 antibody li	1.52e-166
14	262	81.6	426	40	V18558	Mouse 2H7 antibody li	1.52e-166
15	262	81.6	426	40	V18594	Mouse 2H7 antibody li	1.52e-166
16	262	81.6	9208	11	Q65629	Vector contg. TCAE 8	1.52e-166
17	261	81.3	321	6	Q39412	Mab32 light chain V-g	8.07e-166
18	257	80.1	387	4	Q27350	Encodes A5B7 antibody	6.49e-163
19	256	79.8	387	3	Q20983	Encodes Variable regi	3.45e-162
20	254	79.1	732	24	T42508	Murine A5B57 Light ch	9.77e-161
21	252	78.5	435	3	Q15115	IL-2 chimeric antibod	2.76e-159
22	252	78.5	435	6	Q36613	Anti-IL2R-alpha antib	2.76e-159
23	251	78.2	292	29	T63502	Monoclonal antibody A	1.47e-158
24	251	78.2	292	1	Q06213	VK domain of antibody	1.47e-158
25	251	78.2	292	24	T36662	Monoclonal antibody A	1.47e-158
26	246	76.6	420	2	Q11970	Sequence encoding lig	6.22e-155
27	245	76.3	321	4	Q25668	Sequence of the anti-	3.30e-154
28	244	76.0	321	1	N91661	Chimeric antibody lig	1.75e-153
29	244	76.0	639	1	N91657	Chimeric antibody lig	1.75e-153
30	241	75.1	306	29	T63506	Monoclonal antibody C	2.62e-151
31	241	75.1	306	24	T36660	Monoclonal antibody C	2.62e-151
32	241	75.1	726	18	T04024	Anti-EGFR single chai	2.62e-151
33	240	74.8	319	4	Q27174	L-chain variable regi	1.39e-150
34	240	74.8	393	11	Q45427	KM-796 and KM-750 lig	1.39e-150
35	240	74.8	436	1	Q06956	Sequence encoding lig	1.39e-150
36	240	74.8	2169	1	Q06958	Genomic sequence enco	1.39e-150
37	239	74.5	306	1	Q06228	VK domain of antibody	7.37e-150
38	239	74.5	726	18	T04025	Anti-EGFR single chai	7.37e-150
39	238	74.1	408	32	T36307	V region of the L6 li	3.91e-149
40	238	74.1	408	2	N70970	Sequence of the V reg	3.91e-149
41	238	74.1	408	40	V18592	Mouse L6 antibody kap	3.91e-149
42	238	74.1	408	39	V18556	Mouse L6 antibody kap	3.91e-149
43	238	74.1	408	1	N82418	Variable region of th	3.91e-149
44	238	74.1	408	31	T70860	L6 antibody V kappa c	3.91e-149
45	238	74.1	408	30	T51041	Coding sequence for V	3.91e-149

ALIGNMENTS

RESULT 1
ID Q55180 standard; cDNA; 711 BP.
AC Q55180;
DT 21-JUL-1994 (first entry)
DE Sequence encoding the single chain anti-erbB2 antibody, Ab no.23.
KW Single chain anti-erbB1 antibody; cancer therapy; prevention;
KW monoclonal antibody; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT cds 1..171
FT /*tag= a
PN WO9400136-A.
PD 06-JAN-1994.
PF 21-OCT-1992; U08545.
PR 30-JUN-1992; US-906555.
PA (MOLE-) MOLECULAR ONCOLOGY INC.
PI Kasprzyk PG, King CR;
DR WPI; 94-025878/03.
DR P-PSDB; R45442.
PT Treatment of malignancies over-expressing ERB-[2 - using at least
PT 2 monoclonal antibodies which recognise different epitopes on
PT gp185
PS Example; Fig 7; 37pp; English.
CC The source of human erbB-2 protein for the prodn. of antibodies no.
CC 23 and 21 (Q55180/R45442; Q55181/R45443) is a NIH/3T3 cell
CC engineered to express the human erbB-2 protein on its surface (N/
CC erbB-2). Abs no. 21 and 23 are directed against the extracellular
CC domain of gp185 erbB-2. Nude mice manipulated to produce rapidly
CC growing tumours were used in a trial of the efficacy of the Abs. In
CC animals given a combination of the 2 Abs, tumours completely
CC regressed after 11 days.
SQ Sequence 711 BP; 175 A; 183 C; 187 G; 165 T;

Query Match 97.8%; Score 314; DB 9; Length 711;
Best Local Similarity 99.7%; Pred. No. 2.02e-204;
Matches 320; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Db      1 atggacctgcagctgacccagtctccagcaatcctgtctgcatctccaggggagaaggt- 59
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db     60 cacaatgacttgcagggccaccccaagtgtaagttacatgcactggtatcagcagaagcc 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 120

Db    120 aggatcctcccccaaaccttggatttataccacatccaaccctkgcttctggagtccctgc 179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180

Db    180 tcgcttcagtggcggtgggtctgggacctcttactctctcacagtcagcagagtggaggc 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db    240 tgaagatgctgccacttattactgccagcagtggagtcgtagcccacccacgttcggagg 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 300

Db    300 ggggtccaagctggaaataaa 320
          |||||||||||||||||||||
Qy    301 GGGGTCCAAGCTGGAAATAAA 321
```

```
RESULT    2
ID   T65006 standard; cDNA; 711 BP.
AC   T65006;
DT   05-JUN-1997  (first entry)
DE   Single-chain anti-erbB2 antibody e23(Fv) cDNA.
KW   Single chain antibody; variable region; light chain; heavy chain;
KW   breast cancer; ovarian cancer; non-small cell lung carcinoma;
KW   immunodiagnosis; treatment; cytotoxic agent; erbB-2; ds.
OS   Mus musculus.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     1..711
FT                   /*tag= a
FT                   /product= e23(Fv)
PN   US5587458-A.
PD   24-DEC-1996.
PF   07-OCT-1991; 772270.
PR   07-OCT-1991; US-772270.
PR   30-JUN-1992; US-906555.
PR   14-MAY-1993; US-061092.
PA   (ARON-) ARONEX PHARM INC.
PI   Bird RE,  Kasprzyk PG,  King CR;
DR   WPI; 97-064831/06.
DR   P-PSDB; W15185.
PT   Single chain antibodies specific for erbB-2 protein, gp185  - with
PT   labels or cytotoxin, useful for detection and treatment of tumour
PT   cells expressing this protein
PS   Example 8; Columns 25-28; 28pp; English.
CC   The present cDNA sequence codes for a claimed single-chain antibody,
CC   designated e23(Fv), which binds to erbB-2. Monoclonal antibody e23
CC   was generated by immunising mice with N/erbB-2 cells overexpressing
CC   the gp185 protein, removing spleen cells and producing hybridomas
CC   by standard techniques. Messenger RNA coding for the anti-erbB-2
CC   monoclonal antibody was isolated and converted to cDNA. Regions
CC   coding for the heavy- and light- chain variable regions were then
CC   amplified by PCR and joined via a sequence encoding a peptide
CC   linker. The resulting single-chain antibody is useful for in vitro
CC   diagnosis of tumour cells which overexpress the erbB-2 gp185
CC   marker, e.g. breast, ovarian and non-small cell lung carcinomas,
CC   and, when coupled to a cytotoxic agent, to treat such tumours.
SQ   Sequence   711 BP;    175 A;    182 C;    189 G;    164 T;

  Query Match           97.8%;  Score 314;  DB 29;  Length 711;
  Best Local Similarity  99.7%;  Pred. No. 2.02e-204;
  Matches   320;  Conservative    0;  Mismatches   0;  Indels   1;  Gaps   1;

Db      1 atggacctgcagctgacccagtctccagcaatcctgtctgcatctccaggggagaaggt- 59
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db     60 cacaatgacttgcagggccaccccaagtgtaagttacatgcactggtatcagcagaagcc 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 120

Db    120 aggatcctcccccaaaccttggatttataccacatccaaccctkgcttctggagtccctgc 179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180

Db    180 tcgcttcagtggcggtgggtctgggacctcttactctctcacagtcagcagagtggaggc 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db    240 tgaagatgctgccacttattactgccagcagtggagtcgtagcccacccacgttcggagg 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 300

Db    300 ggggtccaagctggaaataaa 320
          |||||||||||||||||||||
Qy    301 GGGGTCCAAGCTGGAAATAAA 321
```

```
RESULT    3
ID   T17728 standard; cDNA; 711 BP.
AC   T17728;
DT   21-MAY-1996  (first entry)
DE   Anti-erbB2 scFv cDNA.
KW   Oncoprotein; erbB2; cell proliferation; tumour; cancer;
KW   intracellular antibody homologue; single chain antibody; scFv;
KW   gene therapy; ds.
OS   Synthetic.
PN   WO9607321-A1.
PD   14-MAR-1996.
PF   23-AUG-1995; U10740.
PR   06-SEP-1994; US-301339.
PR   06-JUN-1995; US-468252.
PA   (UABR-) UAB RES FOUND.
PI   Curiel DT,  Deshane J;
DR   WPI; 96-171307/17.
DR   P-PSDB; R94020.
PT   Inhibition of proliferation or survival of, esp. malignant erbB2,
PT   cells - by introducing nucleic acid mol. encoding antibody homologue
PT   which is expressed and binds, pref. erbB2, protein intracellularly
PS   Claim 42; Page 29-30; 48pp; English.
CC   A nucleic acid comprises a first sequence encoding a signal peptide
CC   (R94019) linked to a second sequence (T17728) encoding a single
CC   chain Fv fragment (R94020) that binds a human erbB2 oncoprotein.
CC   The anti-erbB2 sFv portion is obtained by PCR using e23scFv
CC   plasmid as template.  The signal peptide directs the scFv to the
CC   endoplasmic reticulum.  The nucleic acid is incorporated into a
CC   plasmid or viral vector to facilitate expression of the scFv antibody
CC   homologue within e.g. an epithelial carcinoma cell.  Intracellular
CC   expression of the homologue inhibits surface expression of erbB2 and
CC   thereby inhibits cell proliferation and cell survival and decreases
CC   tumorigenicity.
SQ   Sequence   711 BP;    175 A;    183 C;    189 G;    164 T;

  Query Match           97.2%;  Score 312;  DB 18;  Length 711;
  Best Local Similarity  99.1%;  Pred. No. 5.83e-203;
  Matches   318;  Conservative    1;  Mismatches   1;  Indels   1;  Gaps   1;

Db      1 atggacctgcagctgacccagtctccagcaatcctctctgcatctccaggggagaaggt- 59
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Qy      1 ATGGACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTA 60

Db     60 cacaatgacttgcagggccaccccaagtgtaagttacatgcactggtatcagcagaagcc 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 CACAATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCC 120

Db    120 aggatcctcccccaaaccttggatttataccacatccaacctggcttctggagtccctgc 179
          ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Qy    121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180
```

```
Db    180 tcgcttcagtggcggtgggtctgggacctcttactctctcacagtcagcagagtggaggc 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db    240 tgaagatgctgccacttattactgccagcagtggagtcgtagcccacccacgttcggagg 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGG 300

Db    300 ggggtccaagctggaaataaa 320
          |||||||||||||||||||||
Qy    301 GGGGTCCAAGCTGGAAATAAA 321
```

RESULT 4
ID Q65630 standard; DNA; 384 BP.
AC Q65630;
DT 01-FEB-1995 (first entry)
DE Murine variable region light chain from 2BS.
KW B cell lymphoma chimeric antibody; CD20; peripheral blood cells;
KW cell lysis; ss.
OS Mus musculus.
PN WO9411026-A.
PD 26-MAY-1994.
PF 12-NOV-1993; U10953.
PR 13-NOV-1992; US-978891.
PR 03-NOV-1993; US-149099.
PA (IDEC-) IDEC PHARM CORP.
PI Anderson DR, Hanna N, Leonard JE, Newman RA, Rastetter WH;
PI Reffme;
DR WPI; 94-183162/22.
DR P-PSBD; R55214.
PT Treating B cell lymphoma with chimeric antibody - against CD20,
PT causing rapid depletion of peripheral B cells, also new
PT antibodies and hybridomas
PS Disclosure; Fig 4; 101pp; English.
CC The sequence is the murine variable region light chain derived from
CC murine anti-CD20 monoclonal antibody 2BS.
CC See also Q65629-35.
SQ Sequence 384 BP; 92 A; 106 C; 92 G; 94 T;

Query Match 82.2%; Score 264; DB 11; Length 384;
Best Local Similarity 93.8%; Pred. No. 5.34e-168;
Matches 286; Conservative 1; Mismatches 17; Indels 1; Gaps 1;

```
Db     80 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 138
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db    139 gccagctcaagtgtaagttacatccactggttccagcagaagccaggatcctccccaaa 198
          |||| | ||||||||||||||| |||||||  |||||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    199 ccctggatttatgccacatccaacctggcttctggagtccctgttcgcttcagtggcagt 258
          || ||||||||| |||||||||||||:||||||||||||||| |||||||||||||| ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    259 gggtctgggacttcttactctctcaccatcagcagagtggaggctgaagatgctgccact 318
          |||||||||||| |||||||||||||  ||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    319 tattactgccagcagtggactagtaacccacccacgttcggaggggggaccaagctggaa 378
          |||||||||||||||||| | ||| ||||||||||||||||||||||| ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAA 316

Db    379 atcaa 383
          || ||
Qy    317 ATAAA 321
```

RESULT 5
ID T79900 standard; cDNA; 318 BP.
AC T79900;
DT 27-DEC-1997 (first entry)
DE Anti-Factor IX MAb chimeric light chain cDNA.
KW Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;
KW chimeric antibody; antibody engineering; light chain; ss.
OS Chimeric Mus musculus.
OS Chimeric Homo sapiens.
OS Chimeric synthetic.
PN WO9726010-A1.
PD 24-JUL-1997.
PF 17-JAN-1997; U00759.
PR 24-OCT-1996; US-029119.
PR 17-JAN-1996; US-010108.
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
PI Blackburn MN, Church WR, Feuerstein GZ, Gross MS;
PI Nichols AJ, Padlan EA, Patel AH, Sylvester DR;
DR WPI; 97-385117/35.
DR P-PSDB; W24532.
PT Inhibiting thrombosis with self-limiting antibody to coagulation
PT factor - avoids uncontrolled bleeding by providing only partial
PT inhibition
PS Example 7; Page 128; 150pp; English.
CC This cDNA sequence encodes a mouse-human chimeric antibody
CC light chain (W24532) in which the variable region is derived
CC from mouse anti-human factor IX monoclonal antibody BC2 cDNA (see
CC T79899) and human sequences from the immunoglobulin RF-TS3'CL
CC framework. It was obtained by PCR amplification (see T79897-98)
CC of BC2 cDNA and insertion of the PCR product into F9HZHC 1-3 cDNA
CC (see T77374). Claimed anti-Factor IX chimeric antibodies are
CC useful in the treatment of thrombosis.
SQ Sequence 318 BP; 79 A; 91 C; 78 G; 70 T;

Query Match 81.6%; Score 262; DB 34; Length 318;
Best Local Similarity 93.4%; Pred. No. 1.52e-166;
Matches 285; Conservative 1; Mismatches 18; Indels 1; Gaps 1;

```
Db     14 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 72
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db     73 gccagctcaagtgtaaattacatgcactggtaccagcagaagccaggatcctccccaaa 132
          |||| | |||||||| |||||||||||||| |||||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    133 ccctggatttatgccacatccaacctggcttctggagtccctgctcgcttcagtggcagt 192
          || |||||||| |||||||||||||:||||||||||||||||||||||||||||||| ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    193 gggtctgggacctcttactctctcacaatcagcagagtggaggctgaagatgctgccact 252
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    253 tattactgccagcagtggagtattaacccacggacgttcggtggaggcaccaagctggaa 312
          |||||||||||||||||||  || |||||  |||||||| || ||  ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAA 316

Db    313 atcaa 317
          || ||
Qy    317 ATAAA 321
```

RESULT 6
ID T77377 standard; cDNA; 321 BP.
AC T77377;
DT 26-DEC-1997 (first entry)
DE Mouse anti-human Factor IX antibody BC2 light chain cDNA.
KW Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;
KW humanised antibody; antibody engineering; light chain; CDR;
KW complementarity determining region; myocardial infarction;
KW angina; atrial fibrillation; stroke; kidney damage;
KW pulmonary embolism; deep vein thrombosis; coronary angioplasty;
KW disseminated intravascular coagulation; artificial organ; sepsis;
KW shunt; prosthesis; ss.

```
OS   Mus musculus.
PN   WO9726010-A1.
PD   24-JUL-1997.
PF   17-JAN-1997; U00759.
PR   24-OCT-1996; US-029119.
PR   17-JAN-1996; US-010108.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
PI   Blackburn MN,  Church WR,  Feuerstein GZ,  Gross MS;
PI   Nichols AJ,  Padlan EA, Patel AH,  Sylvester DR;
DR   WPI; 97-385117/35.
DR   P-PSDB; W24520.
PT   Inhibiting thrombosis with self-limiting antibody to coagulation
PT   factor - avoids uncontrolled bleeding by providing only partial
PT   inhibition
PS   Example 5; Page 64; 150pp; English.
CC   This cDNA sequence encodes the light chain variable region (see
CC   W24520) of mouse anti-human Factor IX monoclonal antibody BC2.
CC   Claimed humanised antibodies (see W24510-18) contain CDRs (see
CC   W24504-09) of BC2 heavy and light chains inserted into framework
CC   regions of selected human antibody sequences.  They have self-
CC   limiting neutralising activity, and are useful as anticoagulant
CC   agents in treatment of thrombosis associated with myocardial
CC   infarction, unstable angina, atrial fibrillation, stroke, renal
CC   damage, pulmonary embolism, deep vein thrombosis, percutaneous
CC   transluminal coronary angioplasty, disseminated intravascular
CC   coagulation, sepsis, or artificial organs, shunts or prostheses
CC   (claimed).  Also claimed are chimeric antibodies (see T79900), and
CC   Fab and F(ab')2 fragments.  The claimed antibodies do not cause
CC   uncontrolled bleeding (contrast heparin and warfarin) since they
CC   provide only partial inhibition of coagulation.
SQ   Sequence   321 BP;    78 A;    92 C;    79 G;    72 T;

  Query Match           81.6%;  Score 262;  DB 34;  Length 321;
  Best Local Similarity  93.4%;  Pred. No. 1.52e-166;
  Matches   285;  Conservative    1;  Mismatches  18;  Indels   1;  Gaps   1;

Db     14 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 72
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db     73 gccagctcaagtgtaaattacatgcactggtaccagcagaagccaggatcctcccccaaa 132
           |||| | |||||||||| |||||||||||||| |||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    133 ccctggatttatgccacatccaacctggcttctggagtccctgctcgcttcagtggcagt 192
           || |||||||||  |||||||||||:||||||||||||||||||||||||||||||  ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    193 gggtctgggacctcttactctctcacaatcagcagagtggaggctgaagatgctgccact 252
           ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    253 tattactgccagcagtggagtattaacccacggacgttcggtggaggcaccaagctggaa 312
           ||||||||||||||||||||  || |||||  |||||||| || ||  ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAA 316

Db    313 atcaa 317
           || ||
Qy    317 ATAAA 321


RESULT    7
ID   T79899 standard; cDNA; 335 BP.
AC   T79899;
DT   27-DEC-1997  (first entry)
DE   Anti-Factor IX MAb BC2 light chain PCR product.
KW   Thrombosis; therapy; Factor IX; anticoagulant; monoclonal antibody;
KW   chimeric antibody; antibody engineering; light chain; ss.
OS   Chimeric Mus musculus.
OS   Chimeric synthetic.
PN   WO9726010-A1.
PD   24-JUL-1997.
PF   17-JAN-1997; U00759.
PR   24-OCT-1996; US-029119.
PR   17-JAN-1996; US-010108.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
PI   Blackburn MN,  Church WR,  Feuerstein GZ,  Gross MS;
PI   Nichols AJ,  Padlan EA, Patel AH,  Sylvester DR;
DR   WPI; 97-385117/35.
DR   P-PSDB; W24531.
PT   Inhibiting thrombosis with self-limiting antibody to coagulation
PT   factor - avoids uncontrolled bleeding by providing only partial
PT   inhibition
PS   Example 7; Page 126; 150pp; English.
CC   This cDNA sequence was obtained by PCR amplification (see T79897
CC   and T79898) of the light chain variable region (see also T77377)
CC   of mouse anti-human factor IX monoclonal antibody BC2 cDNA.  The
CC   amplification resulted in the addition of ScaI, NarI ends to the
CC   VL region.  The PCR product was ligated into ScaI, NarI-digested
CC   F9HZHC 1-3 (see T77374) and digested with ScaI, NarI to produce a
CC   mouse-human chimeric light chain F9CHLC (see T79900, W24532).
CC   Claimed anti-Factor IX chimeric antibodies are useful in the
CC   treatment of thrombosis.
SQ   Sequence   335 BP;    80 A;    97 C;    85 G;    73 T;

  Query Match           81.6%;  Score 262;  DB 34;  Length 335;
  Best Local Similarity  93.4%;  Pred. No. 1.52e-166;
  Matches   285;  Conservative    1;  Mismatches  18;  Indels   1;  Gaps   1;

Db     14 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 72
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db     73 gccagctcaagtgtaaattacatgcactggtaccagcagaagccaggatcctcccccaaa 132
           |||| | |||||||||| |||||||||||||| |||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    133 ccctggatttatgccacatccaacctggcttctggagtccctgctcgcttcagtggcagt 192
           || |||||||||  |||||||||||:||||||||||||||||||||||||||||||  ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    193 gggtctgggacctcttactctctcacaatcagcagagtggaggctgaagatgctgccact 252
           ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    253 tattactgccagcagtggagtattaacccacggacgttcggtggaggcaccaagctggaa 312
           ||||||||||||||||||||  || |||||  |||||||| || ||  ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAA 316

Db    313 atcaa 317
           || ||
Qy    317 ATAAA 321


RESULT    8
ID   N70972 standard; cDNA; 403 BP.
AC   N70972;
DT   09-APR-1991  (first entry)
DE   2H7 VL sequence in which the VK gene contains JK5 sequences.
KW   Chimeric antibody; Anti-cancer antibody; ss.
FH   Key             Location/Qualifiers
FT   cds             20..66
FT                   /*tag=  a
FT                   /label=leader peptide
FT   cds             67..138
FT                   /*tag=  b
FT                   /label=FR1
FT   cds             139..165
FT                   /*tag=  c
FT                   /label=CDR1
FT   cds             166..210
FT                   /*tag=  d
FT                   /label=FR2
FT   cds             211..231
```

```
FT                  /*tag=  e
FT                  /label=CDR2
FT   cds            232..327
FT                  /*tag=  f
FT                  /label=FR3
FT   cds            328..354
FT                  /*tag=  g
FT                  /label=CDR3
FT   cds            355..403
FT                  /*tag=  h
FT                  /label=FR4
FT   cds            349..403
FT                  /*tag=  i
FT                  /label=JK5
PN   WO8702671-A.
PD   07-MAY-1987.
PF   27-OCT-1986; U02269.
PR   01-NOV-1985; US-793980.
PA   (ITGE-) INT GENETIC ENG INC.
PA   (ROBI/) ROIBINSON R R.
PI   Robinson RR, Liu AY, Horwitz AH, Wall R;.
DR   WPI; 87-136004/19.
DR   P-PSDB; P70628.
PT   Prodn. of immunoglobulin chains and molecules - is by recombinant
PT   DNA procedures, with chimeric antibodies etc. related to cancer
PT   specific antigens.
PS   Example; Fig 22; 126pp; English.
CC   The patentors claim a chimeric antibody molecule comprising 2 light
CC   chains and 2 heavy chains, each comprising a constant human region
CC   and a variable non-human region.  Coding sequences for the Ig chains
CC   are also claimed.  The invention provides consensus sequences of
CC   light and heavy chain J regions  useful in the design of
CC   oligonucleotides (UIGs) for use as primers or probes for cloning
CC   immunoglobulinlight or heavy chain mRNAs or genes.  Depending on the
CC   nature of design of a particular UIG, it may be capable of
CC   hybridizing to all Ig mRNAs or genes containing a single specific J
CC   sequence. UIG denotes universal immunoglobulin gene.
SQ   Sequence   403 BP;    100 A;    112 C;    93 G;    98 T;

  Query Match           81.6%;  Score 262;  DB 2;  Length 403;
  Best Local Similarity  94.3%;  Pred. No. 1.52e-166;
  Matches   282;  Conservative    1;  Mismatches  15;  Indels   1;  Gaps   1;

Db     99 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 157
           |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db    158 gccagctcaagtgtaagttacatgcactggtaccagcagaagccaggatcctcccccaaa 217
           ||||  | |||||||||||||||||||||||| |||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    218 ccctggatttatgccccatccaacctggcttctggagtccctgctcgcttcagtggcagt 277
           || |||||||||  || ||||||||||:|||||||||||||||||||||||||||||  ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    278 gggtctgggacctcttactctctcacaatcagcagagtggaggctgaagatgctgccact 337
           ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    338 tattactgccagcagtggagttttaacccacccacgttcggtgctgggaccaagctgga 396
           ||||||||||||||||||||  || ||||||||||||||| |  ||| ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGA 315


RESULT    9
ID   N91147 standard; DNA; 403 BP.
AC   N91147;
DT   06-JUL-1990  (first entry)
DE   2H7 Vh sequence.
KW   Antibodies; passive immunisation; pH3-6a; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   cds             20..403
FT                   /*tag=  a
FT   misc_feature    320..352
FT                   /*tag=  b
FT                   /note="JK5 region."
PN   WO8900999-A.
PD   9-FEB-1989.
PF   25-JUL-1988; 02514.
PR   24-JUL-1987; US-077528.
PA   (ITGE-) Int Genetic Eng Inc.
PI   Robinson RR, Liu AY, Horwitz AH, Wall R, Better M;
DR   WPI; 89-061144/08.
DR   P-PSDB; P94778.
PT   Polynucleotide(s) encoding Immunoglobulin molecules -
PT   used for efficient prodn. of chimeric human or non-human or
PT   class switched antibodies.
PS   Disclosure; 7pp; English.
CC   Sequence, derived from M13 subclones of gene fragments, carries the
CC   variable region of chimeric immunoglobulin sequence. The antibodies are
CC   useful in passive immunisation avoiding negative reactions. They are
CC   also useful in assaying and in vitro imaging.
SQ   Sequence   403 BP;   100 A;   112 C;   93 G;   98 T;

  Query Match           81.6%;  Score 262;  DB 1;  Length 403;
  Best Local Similarity  94.3%;  Pred. No. 1.52e-166;
  Matches   282;  Conservative    1;  Mismatches  15;  Indels   1;  Gaps   1;

Db     99 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 157
           |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db    158 gccagctcaagtgtaagttacatgcactggtaccagcagaagccaggatcctcccccaaa 217
           ||||  | |||||||||||||||||||||||| |||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    218 ccctggatttatgccccatccaacctggcttctggagtccctgctcgcttcagtggcagt 277
           || |||||||||  || ||||||||||:|||||||||||||||||||||||||||||  ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    278 gggtctgggacctcttactctctcacaatcagcagagtggaggctgaagatgctgccact 337
           ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    338 tattactgccagcagtggagttttaacccacccacgttcggtgctgggaccaagctgga 396
           ||||||||||||||||||||  || ||||||||||||||| |  ||| ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGA 315


RESULT   10
ID   T51043 standard; cDNA; 426 BP.
AC   T51043;
DT   06-AUG-1997  (first entry)
DE   Coding sequence for light chain variable region of 2H7.
KW   Pectate lyase; signal sequence; Gram-negative bacterium; immunoglobulin;
KW   protein production; human; constant region; passive immunisation; toxin;
KW   serum sickness; anaphylaxis; sweetener; thaumatin; cytoplasm; periplasm;
KW   antibody; Ig; light-chain; hepatitis; mouse; lung carcinoma; cancer;
KW   myeloma cell; B-cell antigen; ss.
OS   Mus musculus.
FH   Key             Location/Qualifiers
FT   cds             43..426
FT                   /*tag=  a
FT                   /product= light chain variable region of 2H7
PN   US5576195-A.
PD   19-NOV-1996.
PF   01-NOV-1985; 793980.
PR   01-NOV-1985; US-793980.
PR   27-OCT-1986; WO-U02269.
PR   24-JUL-1987; US-077528.
PR   11-JAN-1988; US-142039.
PR   29-MAR-1990; US-501092.
PR   08-DEC-1992; US-987555.
PR   22-FEB-1993; US-020671.
PR   09-DEC-1994; US-357234.
```

```
PA   (XOMA ) XOMA CORP.
PI   Better M,  Lei S,  Robinson RR,  Wilcox GL;
DR   WPI; 97-011254/01.
DR   P-PSDB; W10243.
PT   Improved prodn. of protein in Gram -ve bacteria using signal
PT   sequence - from pectate lyase to ensure transport of protein from
PT   the cytoplasm, esp. for prodn. of antibodies
PS   Example 4; Fig 22; 86pp; English.
CC   This sequence represents the coding sequence for the light chain variable
CC   region of the 2H7 mouse monoclonal antibody. The 2H7 antibody recognises
CC   the human B-cell surface antigen Bp35, which plays a role in B-cell
CC   activation. This sequence was used in a human-mouse chimeric antibody
CC   with human B-cell antigen specificity, that was produced using the method
CC   of the invention. The method of the invention is for the production of a
CC   protein in a Gram-negative bacterium. The method improves on current
CC   techniques, by using a vector including DNA encoding the pectate lyase
CC   signal sequence (see T51034), attached to the sequence encoding the
CC   protein for production. The method is especially used to make
CC   immunoglobulins (Ig), particularly those with a human constant region,
CC   suitable for passive immunisation (without risk of serum sickness or
CC   anaphylaxis) or for in vivo/in vitro diagnosis and imaging. The Ig may
CC   also be used therapeutically, optionally coupled to toxins, etc.
CC   Alternatively the protein to be produced is the sweetener thaumatin. The
CC   presence of this signal sequence means that the protein is exported from
CC   the cytoplasm and can be recovered from the culture medium or periplasm,
CC   in active and correctly folded form. The method allows the class of any
CC   antibody to be switched, e.g. most human-human Ig are of M class, easily
CC   reduced and aggregated, and these can now be changed to G, A or E
CC   classes.
SQ   Sequence   426 BP;    100 A;    135 C;    93 G;    98 T;

  Query Match           81.6%;  Score 262;  DB 30;  Length 426;
  Best Local Similarity  94.3%;  Pred. No. 1.52e-166;
  Matches   282;  Conservative    1;  Mismatches  15;  Indels   1;  Gaps   1;

Db    122 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 180
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db    181 gccagctcaagtgtaagttacatgcactggtaccagcagaagccaggatcctcccccaaa 240
          |||| | |||||||||||||||||||||| ||||||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    241 ccctggatttatgccccatccaacctggcttctggagtccctgctcgcttcagtggcagt 300
          || ||||||||| || ||||||||||:|||||||||||||||||||||||||||||||| ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    301 gggtctgggacctcttactctctcacaatcagcagagtggaggctgaagatgctgccact 360
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    361 tattactgccagcagtggagttttaacccacccacgttcggtgctgggaccaagctgga 419
          |||||||||||||||||||  || ||||||||||||| |  ||| ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGA 315


RESULT   11
ID   T36317 standard; cDNA; 426 BP.
AC   T36317;
DT   21-OCT-1997  (first entry)
DE   2H7 antibody light chain variable region cDNA.
KW   Immunoglobulin G; IgG; light chain; recombinant production;
KW   antibody; passive immunisation; serum sickness; anaphylactic shock;
KW   immunoassay; imaging; reagent; complement mediated lysis;
KW   therapy; variable region; ss.
OS   Mus spp.
FH   Key             Location/Qualifiers
FT   signal_peptide  43..108
FT                   /*tag=  a
FT   mat_peptide     109..426
FT                   /*tag=  b
PN   US5595898-A.
PD   21-JAN-1997.
PF   01-NOV-1985; 793980.
PR   29-MAR-1990; US-501092.
PR   01-NOV-1985; US-793980.
PR   27-OCT-1986; WO-U02269.
PR   24-JUL-1987; US-077528.
PR   11-JAN-1988; US-142039.
PR   08-DEC-1992; US-987555.
PR   18-AUG-1994; US-299085.
PA   (XOMA ) XOMA CORP.
PI   Better M,  Horwitz AH,  Lei S,  Liu AY,  Robinson RR;
PI   Wall R,  Wilcox GL;
DR   WPI; 97-107579/10.
DR   P-PSDB; W10589.
PT   Nucleic acid encodinding immunoglobulin fragment - comprising
PT   di:cistronic transcription unit with pectate lyase signal sequences
PS   Example; Fig 22; 95pp; English.
CC   The present sequence encodes the light chain variable region of
CC   the 2H7 antibody. The 2H7 cDNA was used in the preparation of
CC   a novel polynucleotide molecule encoding an Ig fragment. The DNA
CC   molecule comprises 2 DNA sequences encoding 2 pectate lyase
CC   secretion signal sequences respectively linked to a DNA sequence
CC   encoding an Ig Fd molecule or Ig light chain, operably linked to a
CC   single prokaryotic promoter so as to form a dicistronic
CC   transcription unit, provided that the Ig fragment can bind an
CC   antigen and is produced and secreted by an E. coli host cell when
CC   the nucleic acid molecule is expressed in the host cell.
CC   The polynucleotide molecule is used for the production of
CC   recombinant antibodies, which can be used for passive immunisation
CC   without negative immune reactions (e.g. serum sickness and
CC   anaphylactic shock), in labelled forms as immunoassay or imaging
CC   reagents, in complement mediated lysis and for therapeutic
CC   purposes when coupled to a toxin or other therapeutic agent.
SQ   Sequence   426 BP;    100 A;    135 C;    93 G;    98 T;

  Query Match           81.6%;  Score 262;  DB 32;  Length 426;
  Best Local Similarity  94.3%;  Pred. No. 1.52e-166;
  Matches   282;  Conservative    1;  Mismatches  15;  Indels   1;  Gaps   1;

Db    122 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 180
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db    181 gccagctcaagtgtaagttacatgcactggtaccagcagaagccaggatcctcccccaaa 240
          |||| | |||||||||||||||||||||| ||||||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    241 ccctggatttatgccccatccaacctggcttctggagtccctgctcgcttcagtggcagt 300
          || ||||||||| || ||||||||||:|||||||||||||||||||||||||||||||| ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    301 gggtctgggacctcttactctctcacaatcagcagagtggaggctgaagatgctgccact 360
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    361 tattactgccagcagtggagttttaacccacccacgttcggtgctgggaccaagctgga 419
          |||||||||||||||||||  || ||||||||||||| |  ||| ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGA 315


RESULT   12
ID   T70869 standard; cDNA; 426 BP.
AC   T70869;
DT   04-SEP-1997  (first entry)
DE   2H7 light chain variable sequence.
KW   Antibody engineering; heavy chain; light chain; chimaeric antibody;
KW   passive immunisation; diagnosis; hybridoma; monoclonal antibody;
KW   2H7; B-cell antigen; Bp35; ss.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   signal_peptide  43..240
FT                   /*tag=  a
FT   mat_peptide     241..426
FT                   /*tag=  b
```

```
FT   misc_feature    342..375
FT                   /*tag=  c
FT                   /note= "Jk5 sequence element
FT   primer_bind     406..420
FT                   /*tag=  d
FT                   /note= "primer JkHindIII"
PN   US5618920-A.
PD   08-APR-1997.
PF   01-NOV-1985; 793980.
PR   01-NOV-1985; US-793980.
PR   27-OCT-1986; WO-U02269.
PR   24-JUL-1987; US-077528.
PR   11-JAN-1988; US-142039.
PR   29-MAR-1990; US-501092.
PR   17-APR-1992; US-870404.
PR   29-APR-1994; US-235225.
PA   (XOMA ) XOMA CORP.
PI   Better M,  Horwitz AH,  Lei S,  Liu AY,  Robinson RR;
PI   Wall R,  Wilcox GL;
DR   WPI; 97-225473/20.
DR   P-PSDB; W16344.
PT   Secretable immunoglobulin heavy and light chain fragments - capable
PT   of assembling into chimeric antibodies, useful for e.g. passive
PT   immunisation, diagnosis, etc
PS   Example 4; Fig 22; 96pp; English.
CC   A cDNA clone (T70869) codes for the light chain variable region
CC   (W16344) of the 2H7 mouse monoclonal antibody, which recognises
CC   human B-cell surface antigen Bp35.  The sequence was isolated from
CC   a 2H7 cell line cDNA library by PCR amplification.  The 2H7 heavy
CC   chain variable sequence (T60868) has also been isolated.  The
CC   sequences have been used to construct a human-mouse chimaeric
CC   antibody with specificity for the human B-cell antigen.
SQ   Sequence   426 BP;    100 A;    135 C;    93 G;    98 T;

  Query Match           81.6%;  Score 262;  DB 31;  Length 426;
  Best Local Similarity  94.3%;  Pred. No. 1.52e-166;
  Matches   282;  Conservative    1;  Mismatches  15;  Indels   1;  Gaps   1;

Db    122 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 180
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db    181 gccagctcaagtgtaagttacatgcactggtaccagcagaagccaggatcctcccccaaa 240
          |||| | ||||||||||||||||||||||| ||||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    241 ccctggatttatgccccatccaacctggcttctggagtccctgctcgcttcagtggcagt 300
          || ||||||||| || ||||||||||:||||||||||||||||||||||||||||||| ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    301 gggtctgggacctcttactctctcacaatcagcagagtggaggctgaagatgctgccact 360
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    361 tattactgccagcagtggagttttaacccacccacgttcggtgctgggaccaagctgga 419
          |||||||||||||||||||  || ||||||||||||||| |  ||| ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGA 315


RESULT   13
ID   V03927 standard; cDNA; 426 BP.
AC   V03927;
DT   01-JUN-1998  (first entry)
DE   Mouse 2H7 antibody light chain variable region cDNA.
KW   Mouse; murine; light chain; variable region;
KW   immunoglobulin fragment production; Ig fragment production;
KW   monoclonal antibody 2H7; human B-cell surface antigen; ss.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   sig_peptide     43..108
FT                   /*tag=  a
FT   mat_peptide     109..426
FT                   /*tag=  b
PN   US5693493-A.
PD   02-DEC-1997.
PF   25-MAY-1995; 450731.
PR   29-MAR-1990; US-501092.
PR   01-NOV-1985; US-793980.
PR   27-OCT-1986; WO-U02269.
PR   24-JUL-1987; US-077528.
PR   11-JAN-1988; US-142039.
PR   08-DEC-1992; US-987555.
PR   18-AUG-1994; US-299085.
PR   25-MAY-1995; US-450731.
PA   (XOMA ) XOMA CORP.
PI   Better M, Horwitz AH, Lei S, Liu AY, Robinson RR,
PI   Wall R, Wilcox GL;
DR   WPI; 98-031749/03.
DR   P-PSDB; W41071.
PT   Production of chimeric antibody fragments - by culturing E. coli
PT   transformed with dicistronic expression cassette
PS   Example IV; Fig 22; 98pp; English.
CC   The present sequence was used in the development of a novel method
CC   for the production of an immunoglobulin (Ig) fragment capable of
CC   binding an antigen. The method comprises culturing an E. coli host
CC   that has been transformed with a nucleic acid molecule encoding the
CC   Ig fragment, under conditions so that the Ig fragment is produced
CC   and secreted. The nucleic acid molecule comprises DNA sequences
CC   encoding: (a) pectate lyase secretion signal sequence operably
CC   linked to a DNA sequence encoding at least the variable region of
CC   an Ig Fd molecule; and (b) pectate lyase secretion signal sequence
CC   operably linked to a DNA sequence encoding at least the variable
CC   region of an Ig light chain, where (a) and (b) are operably linked
CC   to a single prokaryotic promoter to form a dicistronic
CC   transcription unit. The method is used to produce chimeric Fab
CC   molecules, e.g. derived from murine monoclonal antibody 2H7 raised
CC   against human B-cell surface antigen. The invention provides a
CC   novel approach for producing genetically engineered antibodies of
CC   desired variable region specificity and constant region
CC   properties. The cloned Ig gene products can be produced by
CC   expression in genetically engineered organisms. The application of
CC   chemical gene synthesis, recombinant DNA cloning and production of
CC   specific Ig chains in various organisms provides an effective
CC   solution for the efficient large scale production of human
CC   monoclonal antibodies. The invention also provides a solution to
CC   the problem of class switching antibody molecules.
SQ   Sequence   426 BP;    100 A;    135 C;    93 G;    98 T;

  Query Match           81.6%;  Score 262;  DB 40;  Length 426;
  Best Local Similarity  94.3%;  Pred. No. 1.52e-166;
  Matches   282;  Conservative    1;  Mismatches  15;  Indels   1;  Gaps   1;

Db    122 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 180
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db    181 gccagctcaagtgtaagttacatgcactggtaccagcagaagccaggatcctcccccaaa 240
          |||| | ||||||||||||||||||||||| ||||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    241 ccctggatttatgccccatccaacctggcttctggagtccctgctcgcttcagtggcagt 300
          || ||||||||| || ||||||||||:||||||||||||||||||||||||||||||| ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    301 gggtctgggacctcttactctctcacaatcagcagagtggaggctgaagatgctgccact 360
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    361 tattactgccagcagtggagttttaacccacccacgttcggtgctgggaccaagctgga 419
          |||||||||||||||||||  || ||||||||||||||| |  ||| ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGA 315


RESULT   14
ID   V18558 standard; cDNA; 426 BP.
AC   V18558;
```

```
DT   05-JUN-1998  (first entry)
DE   Mouse 2H7 antibody light chain variable region cDNA.
KW   Mouse; murine; light chain; variable region;
KW   immunoglobulin fragment production; Ig fragment production;
KW   monoclonal antibody 2H7; human B-cell surface antigen; ss.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   sig_peptide     43..108
FT                   /*tag=  a
FT   mat_peptide     109..426
FT                   /*tag=  b
PN   US5698435-A.
PD   16-DEC-1997.
PF   06-JUN-1995; 467140.
PR   29-MAR-1990; US-501092.
PR   01-NOV-1985; US-793980.
PR   27-OCT-1986; WO-U02269.
PR   24-JUL-1987; US-077528.
PR   11-JAN-1988; US-142039.
PR   08-DEC-1992; US-987555.
PR   18-AUG-1994; US-299085.
PR   06-JUN-1995; US-467140.
PA   (XOMA ) XOMA CORP.
PI   Better M, Horwitz AH, Lei S, Liu.AY, Robinson RR,
PI   Wall R, Wilcox GL;
DR   WPI; 98-051492/05.
DR   P-PSDB; W47514.
PT   DNA encoding secretable immunoglobulin fragments - comprising at
PT   least the variable regions of light or heavy chains
PS   Example IV; Fig 22; 98pp; English.
CC   The present sequence was used in the development of a novel method
CC   for the production of an immunoglobulin (Ig) fragment capable of
CC   binding an antigen. The method comprises culturing an E. coli host
CC   that has been transformed with a nucleic acid molecule encoding the
CC   Ig fragment, under conditions so that the Ig fragment is produced
CC   and secreted. The nucleic acid molecule comprises DNA sequences
CC   encoding: (a) pectate lyase secretion signal sequence operably
CC   linked to a DNA sequence encoding at least the variable region of
CC   an Ig Fd molecule; and (b) pectate lyase secretion signal sequence
CC   operably linked to a DNA sequence encoding at least the variable
CC   region of an Ig light chain, where (a) and (b) are operably linked
CC   to a single prokaryotic promoter to form a dicistronic
CC   transcription unit. The method is used to produce chimeric Fab
CC   molecules, e.g. derived from murine monoclonal antibody 2H7 raised
CC   against human B-cell surface antigen. The invention provides a
CC   novel approach for producing genetically engineered antibodies of
CC   desired variable region specificity and constant region
CC   properties. The cloned Ig gene products can be produced by
CC   expression in genetically engineered organisms. The application of
CC   chemical gene synthesis, recombinant DNA cloning and production of
CC   specific Ig chains in various organisms provides an effective
CC   solution for the efficient large scale production of human
CC   monoclonal antibodies. The invention also provides a solution to
CC   the problem of class switching antibody molecules.
SQ   Sequence   426 BP;    100 A;    135 C;    93 G;    98 T;

  Query Match           81.6%;  Score 262;  DB 40;  Length 426;
  Best Local Similarity  94.3%;  Pred. No. 1.52e-166;
  Matches   282;  Conservative    1;  Mismatches  15;  Indels   1;  Gaps   1;

Db    122 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 180
          |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Qy     17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db    181 gccagctcaagtgtaagttacatgcactggtaccagcagaagccaggatcctcccccaaa 240
          |||| | ||||||||||||||||||||||| ||||||||||||||||||||||||||||
Qy     77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db    241 ccctggatttatgccccatccaacctggcttctggagtccctgctcgcttcagtggcagt 300
          || ||||||||| || ||||||||||:||||||||||||||||||||||||||||||| ||
Qy    137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db    301 gggtctgggacctcttactctctcacaatcagcagagtggaggctgaagatgctgccact 360
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Qy    197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db    361 tattactgccagcagtggagttttaacccacccacgttcggtgctgggaccaagctgga 419
          |||||||||||||||||||  || |||||||||||||| |  ||| ||||||||||
Qy    257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGA 315


RESULT   15
ID   V18594 standard; cDNA; 426 BP.
AC   V18594;
DT   05-JUN-1998  (first entry)
DE   Mouse 2H7 antibody light chain variable region cDNA.
KW   Mouse; murine; light chain; variable region;
KW   immunoglobulin fragment production; Ig fragment production;
KW   monoclonal antibody 2H7; human B-cell surface antigen; ss.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   sig_peptide     43..108
FT                   /*tag=  a
FT   mat_peptide     109..426
FT                   /*tag=  b
PN   US5698417-A.
PD   16-DEC-1997.
PF   06-JUN-1995; 466203.
PR   29-MAR-1990; US-501092.
PR   01-NOV-1985; US-793980.
PR   27-OCT-1986; WO-U02269.
PR   24-JUL-1987; US-077528.
PR   11-JAN-1988; US-142039.
PR   08-DEC-1992; US-987555.
PR   18-AUG-1994; US-299085.
PR   25-MAY-1995; US-450731.
PR   06-JUN-1995; US-466203.
PA   (XOMA ) XOMA CORP.
PI   Better M, Horwitz AH, Lei S, Liu AY, Robinson RR,
PI   Wall R, Wilcox GL;
DR   WPI; 98-051487/05.
DR   P-PSDB; W47521.
PT   Production of recombinant immunoglobulin fragment - comprising Fd
PT   molecule and light chain
PS   Example IV; Fig 22; 98pp; English.
CC   The present sequence was used in the development of a novel method
CC   for the production of an immunoglobulin (Ig) fragment capable of
CC   binding an antigen. The method comprises culturing an E. coli host
CC   that has been transformed with a nucleic acid molecule encoding the
CC   Ig fragment, under conditions so that the Ig fragment is produced
CC   and secreted. The nucleic acid molecule comprises DNA sequences
CC   encoding: (a) pectate lyase secretion signal sequence operably
CC   linked to a DNA sequence encoding at least the variable region of
CC   an Ig Fd molecule; and (b) pectate lyase secretion signal sequence
CC   operably linked to a DNA sequence encoding at least the variable
CC   region of an Ig light chain, where (a) and (b) are operably linked
CC   to a single prokaryotic promoter to form a dicistronic
CC   transcription unit. The method is used to produce chimeric Fab
CC   molecules, e.g. derived from murine monoclonal antibody 2H7 raised
CC   against human B-cell surface antigen. The invention provides a
CC   novel approach for producing genetically engineered antibodies of
CC   desired variable region specificity and constant region
CC   properties. The cloned Ig gene products can be produced by
CC   expression in genetically engineered organisms. The application of
CC   chemical gene synthesis, recombinant DNA cloning and production of
CC   specific Ig chains in various organisms provides an effective
CC   solution for the efficient large scale production of human
CC   monoclonal antibodies. The invention also provides a solution to
CC   the problem of class switching antibody molecules.
SQ   Sequence   426 BP;    100 A;    135 C;    93 G;    98 T;

  Query Match           81.6%;  Score 262;  DB 40;  Length 426;
  Best Local Similarity  94.3%;  Pred. No. 1.52e-166;
  Matches   282;  Conservative    1;  Mismatches  15;  Indels   1;  Gaps   1;

Db    122 cccagtctccagcaatcctgtctgcatctccaggggagaaggt-cacaatgacttgcagg 180
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Qy      17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db     181 gccagctcaagtgtaagttacatgcactggtaccagcagaagccaggatcctcccccaaa 240
          |||| | |||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Qy      77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAA 136

Db     241 ccctggatttatgccccatccaacctggcttctggagtccctgctcgcttcagtggcagt 300
          || ||||||||| || ||||||||||:|||||||||||||||||||||||||||||||| ||
Qy     137 CCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGT 196

Db     301 gggtctgggacctcttactctctcacaatcagcagagtggaggctgaagatgctgccact 360
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy     197 GGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACT 256

Db     361 tattactgccagcagtggagttttaacccacccacgttcggtgctgggaccaagctgga 419
          |||||||||||||||||||  || ||||||||||||||| |  ||| ||||||||||
Qy     257 TATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGA 315
```

Search completed: Sat Jan 9 14:32:20 1999
Job time : 65 secs.

```
**************************************************************************
 ___    ___   _____   ____   _____   ____   _    _
|  \  /  |  |  ___|  |  _ \ |  ___| |  _ \ | |  | |
|   \/   |  | |___   | |_| | |___  | |_| | | |__| |
| |\__/| |  |___  |  |  __/ |___  | |\ \  |  __  |
| |    | |   ___| |  | |\ \  ___| | | \ \ | |  | |
|_|    |_|  |_____| |_| \_\ |_____| |_|  \_\|_|  |_|  (TM)
**************************************************************************
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Jan 9 13:35:31 1999; MasPar time 457.62 Seconds
1256.155 Million cell updates/sec

Tabular output not generated.

Title: >US-08-704-178-1
Description: (1-321) from US08704178.seq (2 of 4)
Perfect Score: 321
N.A. Sequence: 1 ATGGACCTGCAGCTGACCCA.....GGGTCCAAGCTGGAAATAAA 321
Comp: TACCTGGACGTCGACTGGGT.....CCCAGGTTCGACCTTTATTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl-est55
1:em_est1 2:em_gss1 3:em_gss2 4:em_gss3
Database: genbank-est107
5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est2 17:gb_est20
18:gb_est21 19:gb_est3 20:gb_est4 21:gb_est5 22:gb_est6
23:gb_est7 24:gb_est8 25:gb_est9 26:gb_gss1 27:gb_gss2
28:gb_gss3 29:gb_gss4

Statistics: Mean 9.931; Variance 1.885; scale 5.267

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	232	72.3	396	11	AA691311	vs14f01.r1 Barstead mo	0.00e+00
2	179	55.8	344	9	AA592800	vo25g11.r1 Barstead mo	0.00e+00
3	91	28.3	345	11	AA710291	vt53a04.r1 Barstead mo	2.27e-131
4	84	26.2	335	24	AA318377	EST20620 Spleen I Homo	3.89e-117
5	81	25.2	393	16	T27593	EST100653 Homo sapiens	4.41e-111
6	80	24.9	488	16	H25625	yl48g05.r1 Homo sapien	4.53e-109
7	79	24.6	194	14	AA921173	vy54b07.r1 Stratagene	4.61e-107
8	79	24.6	333	24	AA300732	EST13847 Testis tumor	4.61e-107
9	79	24.6	413	24	AA301347	EST14279 Testis tumor	4.61e-107
10	78	24.3	253	24	AA295941	EST101165 Thymus III H	4.66e-105
11	78	24.3	370	24	AA295093	EST100400 Pancreas tum	4.66e-105
12	77	24.0	328	24	AA295311	EST100471 Pancreas tum	4.67e-103
13	76	23.7	210	16	R69482	yj83c03.r1 Homo sapien	4.64e-101
14	76	23.7	242	8	AA479857	zu35b05.r1 Soares ovar	4.64e-101
15	76	23.7	276	13	AA869306	vq50c10.r1 Barstead bo	4.64e-101
16	76	23.7	403	25	AA345486	EST51505 Gall bladder	4.64e-101
17	75	23.4	352	24	AA301261	EST14181 Testis tumor	4.57e-99
18	74	23.1	463	16	R69532	yj82d09.r1 Homo sapien	4.46e-97
19	74	23.1	470	19	H62115	yu40h01.r1 Homo sapien	4.46e-97
20	73	22.7	345	25	AA335086	EST39457 Esophagus tum	4.31e-95
21	72	22.4	442	10	AA405415	zu56d02.r1 Soares ovar	4.13e-93
22	70	21.8	328	24	AA318628	EST20808 Spleen I Homo	3.67e-89
23	70	21.8	503	5	R28232	yh57e08.r1 Homo sapien	3.67e-89
24	69	21.5	339	11	AA710249	vt49e01.r1 Barstead mo	3.41e-87
25	69	21.5	395	16	T29112	EST69384 Homo sapiens	3.41e-87
26	69	21.5	467	16	H27034	yl65d05.r1 Homo sapien	3.41e-87
27	68	21.2	438	16	R64693	yi22f09.r1 Homo sapien	3.14e-85
28	67	20.9	287	8	AA496610	zv38g06.r1 Soares ovar	2.85e-83
29	66	20.6	303	24	AA300891	EST14031 Testis tumor	2.56e-81
30	66	20.6	382	24	AA295786	EST100987 Pancreas tum	2.56e-81
31	66	20.6	398	10	AA423447	ve80a03.r1 Soares mous	2.56e-81
32	66	20.6	516	16	R67559	yi42h11.r1 Homo sapien	2.56e-81
33	65	20.2	363	7	AA464313	zx78c12.r1 Soares ovar	2.28e-79
34	65	20.2	363	24	AA300651	EST13764 Testis tumor	2.28e-79
35	64	19.9	334	24	AA327254	EST30547 Colon I Homo	2.00e-77
36	64	19.9	390	16	H45841	yp22b04.r1 Homo sapien	2.00e-77
37	64	19.9	397	5	R49771	yj55c03.r1 Homo sapien	2.00e-77
38	63	19.6	261	5	R46878	yj54h03.r1 Homo sapien	1.73e-75
39	63	19.6	357	25	AA361497	EST71040 T-cell lympho	1.73e-75
40	63	19.6	363	25	AA367405	EST78511 Pancreas tumo	1.73e-75
41	63	19.6	383	16	T29916	EST99871 Homo sapiens	1.73e-75
42	63	19.6	400	7	AA456778	zw27f11.r1 Soares ovar	1.73e-75
43	62	19.3	240	10	AA419046	zv33c09.r1 Soares ovar	1.48e-73
44	62	19.3	248	10	AA430449	zw20f11.r1 Soares ovar	1.48e-73
45	62	19.3	346	24	AA300582	EST13427 Testis tumor	1.48e-73

ALIGNMENTS

```
RESULT    1
LOCUS       AA691311      396 bp    mRNA            EST       16-DEC-1997
DEFINITION  vs14f01.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone 1138201 5' similar to gb:X67211 M.musculus rearranged
            immunoglobulin kappa light chain (MOUSE);, mRNA sequence.
ACCESSION   AA691311
NID         g2692247
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 396)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:619473
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          .1. .396
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
```

```
                polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                from 8 week old mouse.  Colon was harvested 72 hours after
                irradiation with 1400 Gys.  1st strand cDNA was primed
                with  a Not I - oligo(dT) primer
                [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
                T 3'];  double-stranded cDNA was ligated to Eco RI
                adaptors [AATTCGGATCCTTG],  digested with Not I and cloned
                into the Not I and Eco RI sites of  the modified pT7T3
                vector. Library constructed by Bob Barstead."
                /db_xref="taxon:10090"
                /clone="1138201"
                /clone_lib="Barstead mouse irradiated colon MPLRB7"
                /dev_stage="8 weeks"
                /lab_host="DH10B"
BASE COUNT      104 a    107 c     91 g     94 t
ORIGIN

  Query Match           72.3%;  Score 232;  DB 11;  Length 396;
  Best Local Similarity  89.6%;  Pred. No. 0.00e+00;
  Matches   266;  Conservative    1;  Mismatches  29;  Indels   1;  Gaps   1;

Db    101 CTCACCCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAGAAGGT-CACCATGACCTG 159
          || |||||||||||||||||| |||||||||||||||||||||||| ||| ||||| ||
Qy     13 CTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTG 72

Db    160 CAGTGCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGTCAGGCACCTCCCC 219
          ||| |||| | |||||||||||||||||||||||| |||||||| |||| ||||||| 
Qy     73 CAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCC 132

Db    220 CAAAAGATGGATTTATGACACATCCAAACTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGG 279
          ||||   |||||||||  |||||||| ||:|||||||||||||||||||||||||||||
Qy    133 CAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGG 192

Db    280 CAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGCATGGAGGCTGAAGATGCTGC 339
          | |||||||||||||||||||||||||||| |||||||  ||||||||||||||||||
Qy    193 CGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGC 252

Db    340 CACTTATTACTGCCAGCAGTGGAGTAGTAACCCGCTCACGTTCGGTGCTGGGACCAA 396
          |||||||||||||||||||||||||| ||| ||| | ||||||||| |  ||| ||||
Qy    253 CACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAA 309


RESULT    2
LOCUS       AA592800      344 bp    mRNA            EST       17-SEP-1997
DEFINITION  vo25g11.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone 1050980 5' similar to gb:X67211 M.musculus rearranged
            immunoglobulin kappa light chain (MOUSE);, mRNA sequence.
ACCESSION   AA592800
NID         g2406463
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 344)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:582556
            Trace considered overall poor quality
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .344
                      /organism="Mus musculus"
                      /strain="FVB/N"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                      from 8 week old mouse.  Colon was harvested 72 hours after
                      irradiation with 1400 Gys.  1st strand cDNA was primed
                      with  a Not I - oligo(dT) primer
                      [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
                      T 3'];  double-stranded cDNA was ligated to Eco RI
                      adaptors [AATTCGGATCCTTG],  digested with Not I and cloned
                      into the Not I and Eco RI sites of  the modified pT7T3
                      vector. Library constructed by Bob Barstead."
                      /db_xref="taxon:10090"
                      /clone="1050980"
                      /clone_lib="Barstead mouse irradiated colon MPLRB7"
                      /dev_stage="8 weeks"
                      /lab_host="DH10B"
     mRNA             <1. .>344
BASE COUNT      91 a     91 c     70 g     92 t
ORIGIN

  Query Match           55.8%;  Score 179;  DB 9;  Length 344;
  Best Local Similarity  83.7%;  Pred. No. 0.00e+00;
  Matches   226;  Conservative    1;  Mismatches  42;  Indels   1;  Gaps   1;

Db     75 CTCACTCAATCTCCAGCAATCATGTCTGCTTCTCCAGAGGGAGCAATGTCACCACGACCT 134
          || || || |||||||||||| ||||||| ||||||| ||||| |    ||| | ||| |
Qy     13 CTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAG-GGGAGAAGGTACACAATGACTT 71

Db    135 GCAGTGCCAGTTCTACTGTAAGTTACATGCAATGGTAACAGCAGAAGTCAGGTACCTCCC 194
          |||| ||||   | | |||||||||||||||| ||||| |||||||| ||||  ||||||
Qy     72 GCAGGGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCC 131

Db    195 CCAAAACATGGATTTATGATACATCTAAACTGGCTTCTGGAATCCCTGCTCGCTTCAGTG 254
          ||||| | |||||||||   ||||| || ||:||||||||| ||||||||||||||||||
Qy    132 CCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTG 191

Db    255 GCAGTAGCTCTGGGACCTCCTACTCTCTCACAATCAGCAGAATTGAGGCTGAACATGCTG 314
          || || | |||||||||||| |||||||||||| |||||||| | ||||||||| ||||||
Qy    192 GCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTG 251

Db    315 CCACTTATTACTGCCAGCATTGGAGTAGTA 344
          |||||||||||||||||| |||||| |||
Qy    252 CCACTTATTACTGCCAGCAGTGGAGTCGTA 281


RESULT    3
LOCUS       AA710291      345 bp    mRNA            EST       24-DEC-1997
DEFINITION  vt53a04.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone 1166766 5' similar to gb:X02484_rna1 IG KAPPA CHAIN PRECURSOR
            V-IV REGION (HUMAN); gb:J00560 mouse ig kappa mrna from mopc21 &
            other myeloma mrna 3' (MOUSE);, mRNA sequence.
ACCESSION   AA710291
NID         g2720209
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 345)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
```

```
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:632678
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 140.
FEATURES             Location/Qualifiers
     source          1. .345
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                     from 8 week old mouse.  Colon was harvested 72 hours after
                     irradiation with 1400 Gys.  1st strand cDNA was primed
                     with  a Not I - oligo(dT) primer
                     [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     T 3'];  double-stranded cDNA was ligated to Eco RI
                     adaptors [AATTCGGATCCTTG],  digested with Not I and cloned
                     into the Not I and Eco RI sites of  the modified pT7T3
                     vector. Library constructed by Bob Barstead."
                     /db_xref="taxon:10090"
                     /clone="1166766"
                     /clone_lib="Barstead mouse irradiated colon MPLRB7"
                     /dev_stage="8 weeks"
                     /lab_host="DH10B"
BASE COUNT       95 a     81 c     88 g     81 t
ORIGIN

  Query Match           28.3%;  Score 91;  DB 11;  Length 345;
  Best Local Similarity  66.8%;  Pred. No. 2.27e-131;
  Matches   207;  Conservative    1;  Mismatches  99;  Indels   3;  Gaps   3;

Db      35 TGACCCAGTCTCAAAAATTCATGTCCACATCAGCTGGAGACAGGGT-CAGCGTCACCTGC 93
           |||||||||||||  |  | ||  ||||  ||||  | ||  ||  | |||  ||    | ||  |||
Qy      14 TGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGC 73

Db      94 AAGCCAGTCAGAATGTGGGTACTAATGTAGCCTGGTATCAAGAGAAACCAGGGCAATCTC 153
           | | |   |  || |||   | | |  |   |||||||||  ||||  |||||    || |
Qy      74 AGGGCCACCCCAA-GTGTA-AGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCC 131

Db     154 CTAGAGCACTGATTTACTCGGCATCCTACCGGTACAGTGGAGTCCCTGATCGCTTCACAG 213
           | | | |   ||||||  |  ||||| ||| :     ||||||||||| ||||||||  |
Qy     132 CCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTG 191

Db     214 GCAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAATGTGCAGTCTGAAGACTTGG 273
           ||  ||||  ||||||||   || |  |||||||  ||||||  |||  ||  |||||||    |
Qy     192 GCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTG 251

Db     274 CAGAGTATTTCTGTCAGCAATATAACAGCTTTCCGTACACGTTCGGAGGGGGGACCAAGC 333
           |    ||||  |||  |||||  |  |   |    ||   ||||||||||||||||  ||||||
Qy     252 CCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGC 311

Db     334 TGGAAATAAA 343
           ||||||||||
Qy     312 TGGAAATAAA 321


RESULT    4
LOCUS       AA318377      335 bp    mRNA            EST       19-APR-1997
DEFINITION  EST20620 Spleen I Homo sapiens cDNA 5' end similar to similar to
            immunoglobulin kappa light chain, V region, anti-thyroglobulin
            (GB:X79786), mRNA sequence.
ACCESSION   AA318377
NID         g1970863
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 335)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE'    Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .335
                     /organism="Homo sapiens"
                     /note="Organ: spleen; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
                     /db_xref="ATCC (inhost):119058"
                     /db_xref="taxon:9606"
                     /clone_lib="Spleen I"
                     /sex="male"
                     /dev_stage="adult, 23 yrs"
     mRNA            <1. .>335
BASE COUNT       87 a     88 c     71 g     84 t      5 others
ORIGIN

  Query Match           26.2%;  Score 84;  DB 24;  Length 335;
  Best Local Similarity  67.4%;  Pred. No. 3.89e-117;
  Matches   155;  Conservative    1;  Mismatches  74;  Indels   0;  Gaps   0;

Db      36 AGCTATTTACATTGGTATCAGCACAAACCAGGTAAAGCCCCTAAACTCCTGATCTATGNT 95
           ||  ||  | || |||||||||| || |||||     |||| ||||    ||| |||
Qy      92 AGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAACCTTGGATTTATACC 151

Db      96 ACATCCACTTTGCAAAGTGGGGTCCCATCAAGGTTCAGTGGCACTGGATCTGGGACAGAT 155
           |||||||   |:     ||| |||||  |  | |||||||||  ||| ||||||||  |
Qy     152 ACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCT 211

Db     156 TTCACTCTCACCATCAGCAGTCTGCAACCTGAAGATTTTGCAACTTACTACTGTCAACAG 215
           | | |||||||  |||||||  || |  ||||||||  ||| ||||| ||||| || |||
Qy     212 TACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAG 271

Db     216 AGTTCCAGTAGCCCTTGGACGTTCGGCCAAGGGNCCAAGGTGGCAATCAA 265
           |     |||||||    ||||||||    ||| ||||| ||| ||| ||
Qy     272 TGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAAATAAA 321


RESULT    5
```

```
LOCUS       T27593         393 bp    mRNA            EST       06-SEP-1995
DEFINITION  EST100653 Homo sapiens cDNA 5' end similar to immunoglobulin kappa
            light chain, V region (GB:L01279) (HT:3043).
ACCESSION   T27593
NID         g609691
KEYWORDS    EST.
SOURCE      human primer=M13 Reverse library=Human Pancreas.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 393)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W.,
            Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            FitzGerald,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghagen,N.S.M.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle Jr,P.S.,
            Kelley,J.M., Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M.,
            Merrick,J.M., Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T.,
            Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L.,
            Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R.,
            Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A.,
            Coleman,T.A., Collins,E.-J., Dimke,D., Feng,P., Ferrie,A.,
            Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M.,
            Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H.,
            Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C.,
            Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A.,
            Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
  TITLE     Initial Assessment of Human Gene Diversity and Expression Patterns
            Based Upon 52 Million Basepairs of cDNA Sequence
  JOURNAL   Unpublished (1995)
COMMENT     Other_ESTs: THC24356
            Contact: Venter, JC
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: tdbinfo@tdb.tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please contact the TIGR Database
            (tdbinfo@tdb.tigr.org).
FEATURES             Location/Qualifiers
     source          1. .393
                     /organism="Homo sapiens"
     mRNA            <1. .>393
BASE COUNT       88 a    113 c     96 g     94 t      2 others
ORIGIN

  Query Match           25.2%;  Score 81;  DB 16;  Length 393;
  Best Local Similarity  68.3%;  Pred. No. 4.41e-111;
  Matches   185;  Conservative    1;  Mismatches  81;  Indels   4;  Gaps   2;

Db      53 GACATCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACAGAGT-CAC 111
           ||| | ||| ||||||||||||||| |   ||||||||||||  ||| || |  || |||
Qy       4 GACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACAC 63

Db     112 CCTCACTTGCCGGGCAAGTCAGAGCATTAGCACCTTTTTAAATTGGTATCAGCAGAAACC 171
           | |||||| |||| |  |  ||  | ||   | |     | ||||||||||||| ||
Qy      64 AATGACTTGCAGGGCCACCCCAAGTGTAAGTTACATGC---ACTGGTATCAGCAGAAGCC 120

Db     172 AGGGACAGCCCCTAAGCTCCTGATCTCTGCTGCATCCAATTTGCGAAGTGGGGTCCCATC 231
           |||  |  |||| || |    ||| | | |  ||||||| |:     ||| |||||  |
Qy     121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180

Db     232 AAGGTTCAGTGGCTCTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGTCTGCAACC 291
           | |||||||||  ||| ||||||||   || | |||||||  |||||||  || |  |
Qy     181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db     292 TGAAGATGTTGCAACTTACTACTGTCAACAG 322
           |||||||| ||| ||||| ||||| || |||
Qy     241 TGAAGATGCTGCCACTTATTACTGCCAGCAG 271
```

```
RESULT    6
LOCUS       H25625         488 bp    mRNA            EST       10-JUL-1995
DEFINITION  yl48g05.r1 Homo sapiens cDNA clone 161528 5' similar to gb:X72467
            IG KAPPA CHAIN PRECURSOR V-II REGION (HUMAN);.
ACCESSION   H25625
NID         g894748
KEYWORDS    EST.
SOURCE      human clone=161528 library=Soares breast 3NbHBst vector=pT7T3D
            (Pharmacia) with a modified polylinker host=DH10B (ampicillin
            resistant) primer=M13RP1 Rsite1=Not I Rsite2=Eco RI Adult human.
            1st strand cDNA was primed with a Not I - oligo(dT) primer [5'
            TGTTACCAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTT 3'],
            double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia),
            digested with Not I and cloned into the Not I and Eco RI sites of a
            modified pT7T3 vector (Pharmacia). Library went through one round
            of normalization to a Cot = 20. Library constructed by Bento Soares
            and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 488)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 362
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Homo sapiens"
                     /clone="161528"
BASE COUNT       99 a    130 c    128 g    127 t      4 others
ORIGIN

  Query Match           24.9%;  Score 80;  DB 16;  Length 488;
  Best Local Similarity  67.0%;  Pred. No. 4.53e-109;
  Matches   152;  Conservative    1;  Mismatches  74;  Indels   0;  Gaps   0;

Db     168 AGTTATTTGGATTGGTACCTGCAGAAGCCAGGACAGTCTCCACAGCTCCTGATCTATTTG 227
           |||||  || | ||||| | |||||||||||||   || ||  | |    ||| |||
Qy      92 AGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAACCTTGGATTTATACC 151

Db     228 GCTTCTAAGCGGGCCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGAT 287
           | || || | :|| || || |||||||   | |||||||| |||| || || ||   |
Qy     152 ACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCT 211

Db     288 TTTACACTGAGAATCAGTAGAGTGGAGGCTGAGGATGTTGGCTTTTATTACTGCATGCAA 347
           |   | || | | |||| |||||||||||||| |||| || |  ||||||||||  |||
Qy     212 TACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAG 271

Db     348 ACTCTACAGACCCCGCTCACTTTCGGCGGAGGGACCAAGGTNGAGAT 394
           |  | ||| | ||| ||||| || ||| ||||| | || ||
Qy     272 TGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAAAT 318


RESULT    7
```

```
LOCUS       AA921173      194 bp    mRNA            EST       20-APR-1998
DEFINITION  vy54b07.r1 Stratagene mouse lung 937302 Mus musculus cDNA clone
            1299253 5' similar to gb:X67211 M.musculus rearranged
            immunoglobulin kappa light chain (MOUSE);, mRNA sequence.
ACCESSION   AA921173
NID         g3067952
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 194)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:680301
            Trace considered overall poor quality
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .194
                     /organism="Mus musculus"
                     /strain="C57BL/6 x CBA"
                     /note="Organ: lung; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. 6-8 month old female lung and 1.5 year old male
                     lung were source of mRNA. Average insert size: 1.5 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
                     /db_xref="taxon:10090"
                     /clone="1299253"
                     /clone_lib="Stratagene mouse lung 937302"
                     /sex="female"
                     /tissue_type="lung"
                     /dev_stage="6-8 month old"
                     /lab_host="SOLR (kanamycin resistant)"
BASE COUNT       54 a     54 c     39 g     47 t
ORIGIN

  Query Match           24.6%;  Score 79;  DB 14;  Length 194;
  Best Local Similarity  86.8%;  Pred. No. 4.61e-107;
  Matches    99;  Conservative    0;  Mismatches  14;  Indels    1;  Gaps    1;

Db      82 CCCAGTCTCCAGCAATCATGTCTACATCTCCAGGGCAGAAAGT-CACCATGACCTGCAGT 140
           |||||||||||||||||| ||||| |||||||||| |||| || ||| ||||| |||||
Qy      17 CCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAGG 76

Db     141 GCCATCTCAAGTATAAGTAACATGTACTGCTACCAGCAGAAGCCAGGATCCTCC 194
           |||| | ||||| ||||| ||||| |||| || ||||||||||||||||||||
Qy      77 GCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCC 130


RESULT    8
LOCUS       AA300732      333 bp    mRNA            EST       18-APR-1997
DEFINITION  EST13847 Testis tumor Homo sapiens cDNA 5' end similar to similar
            to immunoglobulin kappa light chain (GB:S49006), mRNA sequence.
ACCESSION   AA300732
NID         g1953300
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 333)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: THC87411
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .333
                     /organism="Homo sapiens"
                     /note="Organ: testis; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
                     /db_xref="ATCC (inhost):192041"
                     /db_xref="taxon:9606"
                     /clone_lib="Testis tumor"
                     /sex="male"
                     /dev_stage="adult"
     mRNA            <1. .>333
BASE COUNT       81 a     85 c     80 g     82 t      5 others
ORIGIN

  Query Match           24.6%;  Score 79;  DB 24;  Length 333;
  Best Local Similarity  67.6%;  Pred. No. 4.61e-107;
  Matches   148;  Conservative    1;  Mismatches  70;  Indels    0;  Gaps    0;

Db       3 CTGGTATCAGCAGAAACCAGGGAAAGCCCCTAAGCTCCTGATCTATGCTGCATCCAGTTT 62
           |||||||||||||||| |||||     |||| || |    ||| ||| | ||||||   |
Qy     103 CTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCT 162

Db      63 GCAAAGTGGGGTCCCATCAAGGTTCAGCGGCAGTGGATCTGGGACAGATTTCACTCTCAC 122
           :     ||| |||||  |  | ||||| ||| |||| |||||||||   || | |||||||
Qy     163 KGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCAC 222

Db     123 CATCAGCAGCCTGCAGCCTGAAGATTTTGCAACTTACTATTGTCAACAGGCTAACAGTTT 182
             |||||||  || || ||||||||  ||| ||||| || || || |||   |   ||
Qy     223 AGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAG 282

Db     183 CCCGTGGACGTTCGGCCAAGGGACCAAGGTGGAAATCAA 221
           |||    ||||||||    ||| ||||| ||||||| ||
Qy     283 CCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAAATAAA 321
```

```
RESULT    9
LOCUS       AA301347      413 bp    mRNA            EST       18-APR-1997
DEFINITION  EST14279 Testis tumor Homo sapiens cDNA 5' end similar to similar
            to immunoglobulin kappa light chain, V region (GB:L01279), mRNA
            sequence.
ACCESSION   AA301347
NID         g1953680
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 413)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES              Location/Qualifiers
     source           1. .413
                      /organism="Homo sapiens"
                      /note="Organ: testis; Vector: pBluescript SK-; Site_1:
                      EcoRI; Site_2: XhoI"
                      /db_xref="ATCC (inhost):191658"
                      /db_xref="taxon:9606"
                      /clone_lib="Testis tumor"
                      /sex="male"
                      /dev_stage="adult"
     mRNA             <1. .>413
BASE COUNT      100 a    110 c    105 g     98 t
ORIGIN

  Query Match           24.6%;  Score 79;  DB 24;  Length 413;
  Best Local Similarity  65.8%;  Pred. No. 4.61e-107;
  Matches   212;  Conservative    1;  Mismatches 104;  Indels   5;  Gaps   4;

Db      78 GACATTCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTGGGAGACAGAGT-CAC 136
           ||| | ||| ||||||||||||||| |   |||||||||||||   || || |  || |||
Qy       4 GACCTGCAGCTGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACAC 63

Db     137 CATCACTTGCCGGGCTGGGCAGCACATTAACACCTATTTAAATTGGTATCAACAGAAACC 196
             || |||||| ||||    |  ||  |  | |  ||  |  | ||||||| ||||| ||
Qy      64 AATGACTTGCAGGGCCACCC--CAAGTGTA-AGTTACATGCACTGGTATCAGCAGAAGCC 120

Db     197 AGGGAAAGCCCCTAAGCTCCTGATGTATGGTGCATCCAGTTTGCAAGGTGGGGTCCCTTC 256
           |||     |||| || |    ||| |||    ||||||   |:     ||| |||||| |
Qy     121 AGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGC 180

Db     257 AAGGTTCAGTGGCGGTGGATCTGGGACAGATTTCACTCTCACCATCACCAGTCTGCAACC 316
             | |||||||||||||| ||||||||   || | |||||||  ||| |||  || |  |
Qy     181 TCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGC 240

Db     317 TGAAGATTTTGCAACTTACTACTTTCAGCAGAGTTACAGTCCTCATCAGAAGGTTCGGCC 376
           |||||||  ||| ||||| ||||  |||||| |     ||   |  |  | | |||||
Qy     241 TGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACG-TTCGGAG 299

Db     377 AAGGGACCAAGGTGGAAATCAA 398
           ||| ||||| |||||||  ||
Qy     300 GGGGGTCCAAGCTGGAAATAAA 321


RESULT    10
LOCUS       AA295941      253 bp    mRNA            EST       18-APR-1997
DEFINITION  EST101165 Thymus III Homo sapiens cDNA 5' end similar to similar to
            immunoglobulin kappa light chain, V region (GB:Y00640), mRNA
            sequence.
ACCESSION   AA295941
NID         g1948286
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES              Location/Qualifiers
     source           1. .253
                      /organism="Homo sapiens"
                      /note="Organ: thymus; Vector: pBluescript SK-; Site_1:
                      EcoRI; Site_2: XhoI"
                      /db_xref="ATCC (inhost):191388"
                      /db_xref="taxon:9606"
                      /clone_lib="Thymus III"
```

```
                     /dev_stage="adult"
     mRNA            <1. .>253
BASE COUNT       63 a     69 c     65 g     53 t      3 others
ORIGIN

  Query Match           24.3%;  Score 78;  DB 24;  Length 253;
  Best Local Similarity  66.7%;  Pred. No. 4.66e-105;
  Matches   146;  Conservative    1;  Mismatches  72;  Indels   0;  Gaps   0;

Db      13 CTGGTACCAGCAGAAACCTGGCCAGGNTCCCAGGCTCCTNATCTATGGTGCATCCACCAG 72
           |||||| |||||||| || ||       ||||  |     || |||    |||||| |
Qy     103 CTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAACCTTGGATTTATACCACATCCAACCT 162

Db      73 GGCCACTGGTATCCCAGCCAGGTTCAGTGGCAGTGGGTCTGGGACAGAGTTCACTCTCAC 132
           :||  ||||  |||| ||  | |||||||||| ||||||||||||    | | |||||||
Qy     163 KGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCAC 222

Db     133 CATCAGCAGCCTGCAGTCTGAAGATTTTNCAGTTTATTACTGTCAGCAGTATAAAAACTG 192
           |||||||  || || ||||||||  | |   |||||||||| |||||||  |      |
Qy     223 AGTCAGCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTCGTAG 282

Db     193 GCCAGAGACGTTCGGCCAAGGGACCAAGGTGGAAATCAA 231
           |||   ||||||||    ||| ||||| ||||||| ||
Qy     283 CCCACCCACGTTCGGAGGGGGGTCCAAGCTGGAAATAAA 321


RESULT   11
LOCUS       AA295093      370 bp    mRNA            EST       18-APR-1997
DEFINITION  EST100400 Pancreas tumor I Homo sapiens cDNA 5' end similar to
            similar to immunoglobulin kappa, variable region (GB:Y00640), mRNA
            sequence.
ACCESSION   AA295093
NID         g1947582
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: THC168243
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.

FEATURES             Location/Qualifiers
     source          1. .370
                     /organism="Homo sapiens"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
                     /db_xref="ATCC (inhost):190705"
                     /db_xref="taxon:9606"
                     /clone_lib="Pancreas tumor I"
                     /dev_stage="adult"
     mRNA            <1. .>370
BASE COUNT       85 a    107 c     94 g     77 t      7 others
ORIGIN

  Query Match           24.3%;  Score 78;  DB 24;  Length 370;
  Best Local Similarity  67.3%;  Pred. No. 4.66e-105;
  Matches   175;  Conservative    1;  Mismatches  80;  Indels   4;  Gaps   3;

Db      57 TGACGCANNTNCCAGCCACCCTGTCCGTGTNTCCAGGGGA-AAGAGCCACCCTCTCCTGC 115
           |||| ||    ||||| | |||||| |  | ||||||||| |||   |||  |  | |||
Qy      14 TGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGC 73

Db     116 AGGGCCAGTCAGAGTGTTAGCAGCAACTTAGCCTGGTACCAGCAGAAACCTGGCCAGGCA 175
           |||||||  |  ||||| |  ||  || |   |||||| |||||||| || ||      |
Qy      74 AGGGCCACCCCAAGTGT-A--AGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCC 130

Db     176 CCCAGGCTCCTNATCTATNATGCATCCACCAGGGCCACTGGTATCCCAGCCAGGTTCAGT 235
           ||||  |     || |||    |||||| |  :||  ||||  |||| ||  | ||||||
Qy     131 CCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGT 190

Db     236 GGCAGTGGGTCTGGGACAGAGTTCACTCTCACCATCAGCAGCCTACAGTCTGAAGATTCT 295
           ||| |||||||||||||    | | |||||||  |||||||  |  || ||||||||  ||
Qy     191 GGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCT 250

Db     296 GCAGTTTATTACTGTCAGCA 315
           ||   ||||||||| |||||
Qy     251 GCCACTTATTACTGCCAGCA 270


RESULT   12
LOCUS       AA295311      328 bp    mRNA            EST       18-APR-1997
DEFINITION  EST100471 Pancreas tumor I Homo sapiens cDNA 5' end similar to
            immunoglobulin kappa light chain, VJ regions, mRNA sequence.
ACCESSION   AA295311
NID         g1947646
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 328)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: THC169106
```

```
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES              Location/Qualifiers
     source           1. .328
                      /organism="Homo sapiens"
                      /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                      EcoRI; Site_2: XhoI"
                      /db_xref="ATCC (inhost):190766"
                      /db_xref="taxon:9606"
                      /clone_lib="Pancreas tumor I"
                      /dev_stage="adult"
     mRNA             <1. .>328
BASE COUNT       73 a     93 c     78 g     77 t      7 others
ORIGIN

  Query Match           24.0%;  Score 77;  DB 24;  Length 328;
  Best Local Similarity  66.8%;  Pred. No. 4.67e-103;
  Matches   175;  Conservative    1;  Mismatches  82;  Indels   4;  Gaps   3;

Db      50 TGACGCAGNTTNCAGCCACCCTGTCTGTGTNTCCAGGGGA-AAGAGCCACCCTNTCCTGC 108
           |||| |||  | |||| | |||||||| | ||||||||| |||   |||  |  | |||
Qy      14 TGACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGC 73

Db     109 AGGGCCAGTCAGAGTGTTAGCAGCAACTTAGCCTGGTACCAGCAGAAACCTGGCCAGGCT 168
           |||||||  |  ||||| |  ||  || |   |||||| |||||||| || ||     |
Qy      74 AGGGCCACCCCAAGTGT-A--AGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCC 130

Db     169 CCCAGGCTCCTCATCTATGGTGCATCCATCAGGGNCACTGGCATCCCAGCCAGGTTCAGT 228
           ||||  |     || |||    |||||| |  :|   ||||  |||| ||  | ||||||
Qy     131 CCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGT 190

Db     229 GGCAGTGGGTCTGGGACAGAGTTCACTCTCANCATCAGCATCCTGCAGTCTGAAGATTTT 288
           ||| ||||||||||||||    | | ||||||   ||||||   || || ||||||||  |
Qy     191 GGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCT 250

Db     289 NCAGTTTATTACTGTCAGCAGT 310
           |   ||||||||| |||||||
Qy     251 GCCACTTATTACTGCCAGCAGT 272


RESULT   13
LOCUS       R69482        210 bp    mRNA            EST       01-JUN-1995
DEFINITION  yj83c03.r1 Homo sapiens cDNA clone 155332 5' similar to gb:L09085
            IG KAPPA CHAIN V-I REGION (HUMAN);.
ACCESSION   R69482
NID         g842999
KEYWORDS    EST.
SOURCE      human clone=155332 library=Soares breast 2NbHBst vector=pT7T3D
            (Pharmacia) with a modified polylinker host=DH10B (ampicillin
            resistant) primer=M13RP1 Rsite1=Not I Rsite2=Eco RI Adult female.
            1st strand cDNA was primed with a Not I - oligo(dT) primer [5'
            TGTTACCAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTT 3'],
            double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia),
            digested with Not I and cloned into the Not I and Eco RI sites of a
            modified pT7T3 vector (Pharmacia). Library went through one round
            of normalization to a Cot = 230. Library constructed by Bento
            Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 210)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality.
FEATURES              Location/Qualifiers
     source           1. .210
                      /organism="Homo sapiens"
                      /clone="155332"
BASE COUNT       53 a     56 c     46 g     50 t      5 others
ORIGIN

  Query Match           23.7%;  Score 76;  DB 16;  Length 210;
  Best Local Similarity  68.1%;  Pred. No. 4.64e-101;
  Matches   130;  Conservative    1;  Mismatches  60;  Indels   0;  Gaps   0;

Db       5 CCNAAACTCCTGATCTTTGCTGCATCCACTTTGCAATCAGGGGTCCCATCTCGATTCAGT 64
           || ||||    ||| | | |  ||||||   |:   || || |||||  |||| ||||||
Qy     131 CCCAAACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGT 190

Db      65 GGCAGCNGATCTGGGACAGATTTCACTCTCACCATCAGCAGCCTGCAGCCTGAAGATGTT 124
           ||| |  | ||||||||   || | |||||||  |||||||  || || ||||||||| |
Qy     191 GGCGGTGGGTCTGGGACCTCTTACTCTCTCACAGTCAGCAGAGTGGAGGCTGAAGATGCT 250

Db     125 GCAACTTATTACTGTCAAAAGTACAACANTGNCCCTCACANTTTCGGCGGAGGGACCAAG 184
           || |||||||||| ||  |||  |    |  ||| | ||  ||||| || ||| |||||
Qy     251 GCCACTTATTACTGCCAGCAGTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAG 310

Db     185 GTAGAGATCAA 195
           | || || ||
Qy     311 CTGGAAATAAA 321


RESULT   14
LOCUS       AA479857      242 bp    mRNA            EST       08-AUG-1997
DEFINITION  zu35b05.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 739953
            5' similar to gb:X06764 IG KAPPA CHAIN PRECURSOR V-III REGION
            (HUMAN);, mRNA sequence.
ACCESSION   AA479857
NID         g2205743
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 242)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
            Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
```

```
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES              Location/Qualifiers
     source           1. .242
                      /organism="Homo sapiens"
                      /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library constructed by Bento Soares and
                      M.Fatima Bonaldo."
                      /db_xref="taxon:9606"
                      /clone="739953"
                      /clone_lib="Soares ovary tumor NbHOT"
                      /sex="Female"
                      /tissue_type="ovarian tumor"
                      /lab_host="DH10B (ampicillin resistant)"
     mRNA             <1. .>242
                      /db_xref="GDB:5939639"
BASE COUNT       54 a     66 c     69 g     53 t
ORIGIN

  Query Match           23.7%;  Score 76;  DB 8;  Length 242;
  Best Local Similarity  67.6%;  Pred. No. 4.64e-101;
  Matches   152;  Conservative    1;  Mismatches  71;  Indels   1;  Gaps   1;

Db      18 AGCTACTTAGCCTGGTACCAGCAGAAACCTGGCCAGGCTCCCAGGCTCCTCATCTATGGT 77
           || ||| |   |||||| |||||||| || ||     | ||||  |     || |||
Qy      92 AGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAAACCTTGGATTTATACC 151

Db      78 GCATCCAGCAGGGCCACTGGCATCCCAGACAGGTTCAGTGGCAGTGGGTCTGGGACAGAC 137
            |||||| |  :||  ||||  |||| |   | |||||||| ||||||||||||
Qy     152 ACATCCAACCTKGCTTCTGGAGTCCCTGCTCGCTTCAGTGGCGGTGGGTCTGGGACCTCT 211

Db     138 TTCACTCTCACCATCATGCAGACTGGAGCCTGAAGATTTTGCAGTGTATTACTGTCAGCA 197
           | | |||||||  ||| ||||| ||||| ||||||||  |||   |||||||| |||||
Qy     212 TACTCTCTCACAGTCA-GCAGAGTGGAGGCTGAAGATGCTGCCACTTATTACTGCCAGCA 270

Db     198 GTATGGTAGCTCACCGCTCACTTTCGGCGGAGGGACCAAGGTGGA 242
           ||   || |    || | ||| |||||  || ||| |||||  ||||
Qy     271 GTGGAGTCGTAGCCCACCCACGTTCGGAGGGGGGTCCAAGCTGGA 315


RESULT   15
LOCUS       AA869306      276 bp    mRNA            EST       16-MAR-1998
DEFINITION  vq50c10.r1 Barstead bowel MPLRB9 Mus musculus cDNA clone 1097682 5'
            similar to gb:X67211 M.musculus rearranged immunoglobulin kappa
            light chain (MOUSE);, mRNA sequence.
ACCESSION   AA869306
NID         g2964751
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 276)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:603914
            Putative full length read
            vector to vector length is 724
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 242.
FEATURES              Location/Qualifiers
     source           1. .276
                      /organism="Mus musculus"
                      /strain="FVB/N"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      [AATTCGGATCCTTG], digested with Not I and cloned into the
                      Not I and Eco RI sites of the modified pT7T3 vector.
                      Source irradiated bowel harvested 72 hours after
                      irradiation (1400 Gys). Library constructed by Bob
                      Barstead."
                      /db_xref="taxon:10090"
                      /clone="1097682"
                      /clone_lib="Barstead bowel MPLRB9"
                      /tissue_type="bowel"
                      /dev_stage="8 weeks"
                      /lab_host="DH10B"
BASE COUNT       83 a     63 c     59 g     71 t
ORIGIN

  Query Match           23.7%;  Score 76;  DB 13;  Length 276;
  Best Local Similarity  83.9%;  Pred. No. 4.64e-101;
  Matches   135;  Conservative    1;  Mismatches  18;  Indels   7;  Gaps   7;

Db     123 ACCCAGTCTC-AGCAATCATGTCTGCATCTCTAGCGGAGAAGGT-CACA-TGAGCTGCAG 179
           |||||||||| ||||||| |||||||||||| || ||||||||| |||| |||  |||||
Qy      16 ACCCAGTCTCCAGCAATCCTGTCTGCATCTCCAGGGGAGAAGGTACACAATGACTTGCAG 75

Db     180 GGC-AGCTCAAGTGTAAATTACATGTACTGGTACCAGCAGAAGTCAGATGCCTCCCC-AA 237
           ||| | | |||||||||| ||||||| ||||||| ||||||||| |||   ||||||| ||
Qy      76 GGCCACCCCAAGTGTAAGTTACATGCACTGGTATCAGCAGAAGCCAGGATCCTCCCCCAA 135

Db     238 ACTATGGATTTATTACACATCCAAC-TGGCT-CTGGAGTCC 276
           ||  |||||||||  |||||||||| |:||| |||||||||
Qy     136 ACCTTGGATTTATACCACATCCAACCTKGCTTCTGGAGTCC 176


Search completed: Sat Jan  9 13:43:17 1999
Job time : 466 secs.
```

```
**************************************************************************
      ___    ___   _______   _____   _______   _____   _     _
     |   \  /   | |   ____| |  __ \ |   ____| |  __ \ | |   | |
     | |\ \/ /| | |  |___  | |__| | |  |___  | |__| | |___| |
     | | \__/ | | |   ___| |  ___/ |   ___| |  ___/ |  ___  |
     | |      | | |  |____ | |\ \   |  |____ | |\ \  | |   | |
     |_|      |_| |_______| |_| \_\  |_______| |_| \_\ |_|   |_|  (TM)
**************************************************************************

        Release 3.1A John F. Collins, Biocomputing Research Unit.
        Copyright (c) 1993-1998  University of Edinburgh, U.K.
              Distribution rights by Oxford Molecular Ltd

MPsrch_nn   n.a. -  n.a. database search, using Smith-Waterman algorithm

Run on:          Sat Jan  9 15:15:33 1999;  MasPar time 533.48 Seconds
                                            1437.582 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-704-178-1
Description:     (364-711) from US08704178.seq (3 of 4)
Perfect Score:   348
N.A. Sequence:      364 TGTGCAGCTGCAGGAGTCAG.........GGGACCACGGTCACCGTCTC 711
     Comp:              ACACGTCGACGTCCTCAGTC.........CCCTGGTGCCAGTGGCAGAG

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        567134 seqs, 1101898692 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        embl55
                   1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
                   7:em_om 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
                   13:em_ro 14:em_vi
Database:        genbank107
                   15:gb_ba1 16:gb_ba2 17:gb_htg 18:gb_in 19:gb_om 20:gb_ov
                   21:gb_pat 22:gb_ph 23:gb_pl1 24:gb_pl2 25:gb_pr1
                   26:gb_pr2 27:gb_pr3 28:gb_ro 29:gb_st 30:gb_sts 31:gb_sy
                   32:gb_un 33:gb_vi

Statistics:      Mean 9.880;  Variance 4.986;  scale 1.981

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	348	100.0	711	21	I32406	Sequence 1 from patent	6.18e-258
2	348	100.0	711	31	S39590	anti-erbB2 imnmunotoxi	6.18e-258
3	282	81.0	336	28	MMU37860	Mus musculus rearrange	7.81e-203
4	265	76.1	354	28	S78745	Ig VH=anti-phosphatidy	1.05e-188
5	264	75.9	351	28	MMU25678	Mus musculus Ig heavy	7.10e-188
6	262	75.3	354	28	MMAHCVR2	M.musculus antibody he	3.24e-186
7	259	74.4	355	28	S74055	IgG VH IIA=IgG2b VH re	1.00e-183
8	259	74.4	357	28	MUSCMVG	Mouse mRNA sequence, p	1.00e-183
9	259	74.4	414	21	I78572	Sequence 84 from paten	1.00e-183
10	259	74.4	414	21	I78627	Sequence 84 from paten	1.00e-183
11	259	74.4	414	21	I31960	Sequence 84 from paten	1.00e-183
12	258	74.1	421	28	S51851	M104E myeloma immunogl	6.75e-183
13	257	73.9	723	28	AF035617	Mus musculus scFv anti	4.56e-182
14	252	72.4	366	28	AF081541	Mus musculus anti-HIV-	6.36e-178
15	250	71.8	440	28	MUSHA4	Mouse active H-chain V	2.88e-176
16	249	71.6	411	28	MMU88682	Mus musculus anti-DNA	1.94e-175
17	249	71.6	443	21	E07912	cDNA encoding H chain	1.94e-175
18	248	71.3	383	28	MMU28805	Mus musculus MoAb Mc3.	1.31e-174
19	248	71.3	428	28	MUSIGGVDJE	Mouse immunoglobulin h	1.31e-174
20	247	71.0	357	28	AF004402	Mus musculus anti-DNA	8.80e-174
21	246	70.7	306	28	MUSIGHI	Mouse IgH chain VDJ4-r	5.92e-173
22	246	70.7	440	28	MUSHA3	Mouse active H-chain V	5.92e-173
23	246	70.7	462	28	MUSIGKCLS	Mouse IgMk rearranged	5.92e-173
24	245	70.4	408	28	MUSIGHAB	Mouse rearranged heavy	3.98e-172
25	243	69.8	354	28	MDIGMVAE	M.domesticus IgM varia	1.80e-170
26	243	69.8	366	28	AF083188	Mus musculus anti-BSA	1.80e-170
27	243	69.8	424	28	MUSIGHCP	Mouse rearranged heavy	1.80e-170
28	242	69.5	366	28	MMVAR604	m.musculus mRNA for im	1.21e-169
29	240	69.0	294	28	MUSIGHAAV	Mouse Ig gamma-chain (	5.45e-168
30	240	69.0	346	28	MMIGHT110	Mouse immunoglobulin v	5.45e-168
31	240	69.0	393	28	MUSIGGVDJB	Mouse immunoglobulin h	5.45e-168
32	240	69.0	916	21	I45604	Sequence 29 from paten	5.45e-168
33	239	68.7	345	28	MMVAR609	m.musculus mRNA for im	3.66e-167
34	239	68.7	363	28	MMU60457	Mus musculus Ig anti-D	3.66e-167
35	238	68.4	339	28	MUSIG438A	Mouse rearranged immun	2.46e-166
36	238	68.4	351	28	MMVH343C	Mouse mRNA for immunog	2.46e-166
37	238	68.4	351	28	MUSANTDNAS	Mouse anti-DNA antibod	2.46e-166
38	238	68.4	354	28	MMBV1619H	M.musculus heavy chain	2.46e-166
39	238	68.4	547	28	MUSIGHAPA	Mus musculus DNA for I	2.46e-166
40	237	68.1	339	28	MMIGPE15	M.musculus mRNA for im	1.65e-165
41	237	68.1	351	28	S54755	IgG3 VH=anti-DNA IgG3	1.65e-165
42	237	68.1	351	28	MMAFBPOVH	M.musculus mRNA for an	1.65e-165
43	236	67.8	2911	28	MUSIGHZZH	Mouse Ig germline heav	1.11e-164
44	235	67.5	354	21	E07407	cDNA encoding H chain	7.41e-164
45	235	67.5	354	28	S75677	Ig VH3=anti-human glyc	7.41e-164

ALIGNMENTS

```
RESULT    1
LOCUS       I32406         711 bp    DNA             PAT       07-JAN-1997
DEFINITION  Sequence 1 from patent US 5587458.
ACCESSION   I32406
NID         g1823197
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 711)
  AUTHORS   King,C.Richter, Kasprzyk,P.G. and Bird,R.E.
  TITLE     Anti-erbB-2 antibodies, combinations thereof, and therapeutic and
            diagnostic uses thereof
  JOURNAL   Patent: US 5587458-A 1 24-DEC-1996;
FEATURES             Location/Qualifiers
     source          1. .711
                     /organism="unknown"
BASE COUNT      175 a    182 c    190 g    164 t
ORIGIN

  Query Match          100.0%;  Score 348;  DB 21;  Length 711;
  Best Local Similarity 100.0%;  Pred. No. 6.18e-258;
  Matches   348;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    363 TGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATC 422
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    364 TGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATC 423

Db    423 CTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCA 482
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    424 CTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCA 483

Db    483 TGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAA 542
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    484 TGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAA 543

Db    543 CCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACAT 602
```

```
Qy    544 CCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACAT 603

Db    603 GGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTAC 662
Qy    604 GGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTAC 663

Db    663 GGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 710
Qy    664 GGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711
```

```
RESULT    2
LOCUS       S39590        711 bp                    SYN       10-FEB-1993
DEFINITION  anti-erbB2 imnmunotoxin antigen binding region [mice, Other
            Synthetic Recombinant Partial, 711 nt].
ACCESSION   S39590
NID         g251113
KEYWORDS    .
SOURCE      Mus sp.
  ORGANISM  Mus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 711)
  AUTHORS   Batra,J.K., Kasprzyk,P.G., Bird,R.E., Pastan,I. and King,C.R.
  TITLE     Recombinant anti-erbB2 immunotoxins containing Pseudomonas exotoxin
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 89 (13), 5867-5871 (1992)
  MEDLINE   92335198
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 108547] from the original journal article.
            This sequence comes from Fig. 1.
FEATURES             Location/Qualifiers
     source          1. .711
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
     gene            1. .711
                     /partial
                     /gene="anti-erbB2 imnmunotoxin antigen binding region"
     CDS             1. .711
                     /partial
                     /gene="anti-erbB2 imnmunotoxin antigen binding region"
                     /note="This sequence comes from Fig. 1."
                     /codon_start=1
                     /db_xref="PID:g251114"
                     /translation="MDLQLTQSPAILSASPGEKVTMTCRATPSVSYMHWYQQKPGSSP
                     KPWIYTTSNLASGVPARFSGGGSGTSYSLTVSRVEAEDAATYYCQQWSRSPPTFGGGS
                     KLEIKGSTSGSGKSSEGKGVQLQESGPEVVKPGGSMKISCKTSGYSFTGHTMNWVKQS
                     HGKNLEWIGLINPYNGDTNYNQKFKGKATFTVDKSSSTAYMELLSLTSEDSAVYYCAR
                     RVTDWYFDVWGAGTTVTVS"
BASE COUNT      175 a    182 c    190 g    164 t
ORIGIN
```

```
  Query Match          100.0%;  Score 348;  DB 31;  Length 711;
  Best Local Similarity 100.0%;  Pred. No. 6.18e-258;
  Matches   348;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    363 TGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATC 422
Qy    364 TGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATC 423

Db    423 CTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCA 482
Qy    424 CTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCA 483

Db    483 TGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAA 542
Qy    484 TGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAA 543

Db    543 CCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACAT 602
Qy    544 CCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACAT 603

Db    603 GGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTAC 662
Qy    604 GGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTAC 663

Db    663 GGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 710
Qy    664 GGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711
```

```
RESULT    3
LOCUS       MMU37860      336 bp    mRNA            ROD       16-MAR-1996
DEFINITION  Mus musculus rearranged immunoglobulin heavy chain variable region
            mRNA, partial cds.
ACCESSION   U37860
NID         g1127634
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 336)
  AUTHORS   Caton,A.J., Swartzentruber,J.R., Kuhl,A.L., Carding,S.R. and
            Stark,S.E.
  TITLE     Activation and negative selection of functionally distinct subsets
            of antibody-secreting cells by influenza hemagglutinin as a viral
            and a neo-self antigen
  JOURNAL   J. Exp. Med. 183 (1), 13-26 (1996)
  MEDLINE   96136744
REFERENCE   2  (bases 1 to 336)
  AUTHORS   Swartzentruber,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-OCT-1995) Jennifer R. Swartzentruber, The Wistar
            Institute, 3601 Spruce St., Philadelphia, PA 19104, USA
FEATURES             Location/Qualifiers
     source          1. .336
                     /organism="Mus musculus"
                     /isolate="h922-163"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
                     /cell_type="B-lymphocyte"
     CDS             <1. .>336
                     /codon_start=1
                     /product="rearranged immunoglobulin heavy chain variable
                     region"
                     /db_xref="PID:g1127635"
                     /translation="ELVKPGASMKISCKASGYSFTGYTMNWVKQSHGKNLEWIGLINP
                     YNGGTSYNQKFKGKATLTVDKSSSTAYMELLSLTSEDSAVYYCAREWLLRYFDVWGAG
                     TTVTVSSAKT"
BASE COUNT       92 a     81 c     85 g     78 t
ORIGIN
```

```
  Query Match          81.0%;  Score 282;  DB 28;  Length 336;
  Best Local Similarity 95.1%;  Pred. No. 7.81e-203;
  Matches   308;  Conservative    0;  Mismatches  14;  Indels   2;  Gaps   2;

Db      1 GAGCTGGTGAAGCCTGGAGCTTCAATGAAGATATCCTGCAAGGCTTCTGGTTACTCATTC 60
Qy    389 GAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCCTGCAAGACTTCTGGTTACTCATTC 448

Db     61 ACTGGCTACACCATGAACTGGGTGAAGCAGAGCCATGGAAAGAACCTTGAGTGGATTGGA 120
Qy    449 ACTGGCCACACCATGAACTGGGTGAAGCAGAGCCATGGAAAGAACCTTGAGTGGATTGGA 508

Db    121 CTTATTAATCCTTACAATGGTGGTACTAGCTACAACCAGAAGTTCAAGGGCAAGGCCACA 180
Qy    509 CTTATTAATCCTTACAATGGTGATACTAACTACAACCAGAAGTTCAAGGGCAAGGCCACA 568

Db    181 TTAACTGTAGACAAGTCATCCAGCACAGCCTACATGGAGCTCCTCAGTCTGACATCTGAG 240
Qy    569 TTTACTGTAGACAAGTCGTCCAGCACAGCCTACATGGAGCTCCTCAGTCTGACATCTGAG 628

Db    241 GACTCTGCAGTCTATTACTGTGCAAGAGAGTGGTTACTAC-GGTACTTCGATGTCTGGGG 299
```

```
          |||||||||||||||||||||||||| ||| | |    || ||||||||||||||||||
Qy    629 GACTCTGCAGTCTATTACTGTGCAAG-GAGGGTTACGGACTGGTACTTCGATGTCTGGGG 687

Db    300 CGCAGGGACCACGGTCACCGTCTC 323
          |||||||||||||||||||||||||
Qy    688 CGCAGGGACCACGGTCACCGTCTC 711


RESULT    4
LOCUS       S78745        354 bp    mRNA            ROD       28-OCT-1995
DEFINITION  Ig VH=anti-phosphatidylserine monclonal antibody PS4A7 heavy chain
            variable region [mice, hybridoma, mRNA Partial, 354 nt].
ACCESSION   S78745
NID         g1041920
KEYWORDS    .
SOURCE      Mus sp. hybridoma.
  ORGANISM  Mus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 354)
  AUTHORS   Igarashi,K., Asai,K., Kaneda,M., Umeda,M. and Inoue,K.
  TITLE     Specific binding of a synthetic peptide derived from an antibody
            complementarity determining region to phosphatidylserine
  JOURNAL   J. Biochem. 117 (2), 452-457 (1995)
  MEDLINE   95332275
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 169252] from the original journal article.
FEATURES             Location/Qualifiers
     source          1. .354
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
     gene            1. .354
                     /partial
                     /note="anti-phosphatidylserine monclonal antibody PS4A7
                     heavy chain variable region"
                     /gene="Ig VH"
     CDS             1. .354
                     /partial
                     /gene="Ig VH"
                     /note="This sequence comes from Fig. 1A.  Protein sequence
                     is in conflict with the conceptual translation;
                     mismatch(35[K->N])"
                     /codon_start=1
                     /product="anti-phosphatidylserine monclonal antibody PS4A7
                     heavy chain variable region"
                     /db_xref="PID:g1041921"
                     /translation="EVQLQQSGPELVKPGASMKISCKASGYSFTGYTMKWVKQSHGKN
                     LEWIGLINPYNGGTSYNQKFKGKATLTVDKSSSTAYMELLSLTSDDSAVYYCAREGDY
                     DGAMDYWGQGTSVTVS"
BASE COUNT       96 a     84 c     92 g     82 t
ORIGIN

  Query Match           76.1%;  Score 265;  DB 28;  Length 354;
  Best Local Similarity  90.0%;  Pred. No. 1.05e-188;
  Matches   315;  Conservative    0;  Mismatches  32;  Indels   3;  Gaps   2;

Db      4 GTCCAGCTGCAACAGTCTGGACCTGAGCTGGTGAAGCCTGGAGCTTCAATGAAGATATCC 63
          || ||||||||  |||| |||||||||| ||||||||||||||| ||||||||||||||||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     64 TGCAAGGCTTCTGGTTACTCATTCACTGGCTACACCATGAACTGGGTGAAGCAGAGCCAT 123
          |||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    124 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGGTACTAGCTACAAC 183
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    184 CAGAAGTTCAAGGGCAAGGCCACATTAACTGTAGACAAGTCTTCCAGCACAGCCTACATG 243
          |||||||||||||||||||||||||| ||||||||||||||| ||||||||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    244 GAGCTCCTCAGTCTGACATCTGATGACTCTGCAGTCTATTACTGTGCAAGAGAAGGGGAT 303
          |||||||||||||||||||||||| ||||||||||||||||||||||||||  |||  |
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG-GA--GGGTT 661

Db    304 TACGACGGGGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTC 353
             ||| ||    | ||   ||||||   ||| ||| | ||||||||||
Qy    662 ACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT    5
LOCUS       MMU25678      351 bp    mRNA            ROD       08-SEP-1998
DEFINITION  Mus musculus Ig heavy chain variable region mRNA, partial cds.
ACCESSION   U25678
NID         g847823
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 351)
  AUTHORS   Liu,J., Wu,W., Wang,B., Pan,Z., Tian,P., Huang,Z. and Hong,J.
  TITLE     Cloning and sequencing of immunoglobulin heavy variable-region gene
            by polymerase chain reaction
  JOURNAL   Sheng Wu Hua Hsueh Tsa Chih 9 (No.1), 70-75 (1993)
REFERENCE   2  (bases 1 to 351)
  AUTHORS   Xu,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-APR-1995) Lin Xu, Institute of Biophysics, Academia
            Sinica, Dept. of Protein Engineering, 15 Datun Road, Chaoyang
            District, Beijing 100101, Peoples Republic of China
FEATURES             Location/Qualifiers
     source          1. .351
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             <1. .>351
                     /codon_start=1
                     /label=JC1077
                     /product="immunoglobulin heavy chain variable region"
                     /db_xref="PID:g847824"
                     /translation="QVQLQQSGPELVKPGASMKISCKTSGYSFTGYTMNWVKQSHEKN
                     LEWIGLINPYSGGTNYNQKFKGKATLTVDKSSSTAYMELLSLTSEDSAVYYCARYYSK
                     AMDYWGQGTTVTVSS"
BASE COUNT       98 a     88 c     84 g     81 t
ORIGIN

  Query Match           75.9%;  Score 264;  DB 28;  Length 351;
  Best Local Similarity  90.2%;  Pred. No. 7.10e-188;
  Matches   313;  Conservative    0;  Mismatches  31;  Indels   3;  Gaps   2;

Db      4 GTCCAACTGCAGCAGTCAGGGCCTGAGCTGGTGAAACCTGGTGCTTCAATGAAGATATCC 63
          || || |||||| ||||||| |||||| ||||||| ||||| | ||||||||||||||||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     64 TGCAAGACTTCTGGTTACTCATTCACTGGCTACACCATGAACTGGGTGAAGCAGAGCCAT 123
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    124 GAAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAGTGGTGGTACTAACTACAAC 183
          | |||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    184 CAGAAGTTCAAGGGCAAGGCCACATTAACTGTAGACAAGTCATCCAGCACAGCCTACATG 243
          |||||||||||||||||||||||||| ||||||||||||| |||||||||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    244 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG-ATACTACTC 302
          |||||||||||||||||||||||||||||||||||||||||||||||||| |   |
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTACG 664

Db    303 TAAGGCTA--TGGACTACTGGGGCCAAGGGACCACGGTCACCGTCTC 347
            |  | ||  | ||   |||||||  ||||||||||||||||||||
Qy    665 GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711
```

```
RESULT    6
LOCUS       MMAHCVR2      354 bp    DNA             ROD       07-OCT-1996
DEFINITION  M.musculus antibody heavy chain variable region (354bp).
ACCESSION   X90877
NID         g1518294
KEYWORDS    antibody heavy chain; variable region.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 354)
  AUTHORS   Geiser,M. and Kretzschmar,T.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 354)
  AUTHORS   Geiser,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-1995) M. Geiser, CIBA-GEIGY LTD, CDDT,
            K-681.5.46, CH-4002, Basel, SWITZERLAND
FEATURES             Location/Qualifiers
     source          1. .354
                     /organism="Mus musculus"
                     /strain="Balb/c"
                     /db_xref="taxon:10090"
                     /dev_stage="12 weeks old"
                     /tissue_type="spleen"
                     /cell_type="lymphocytes"
     CDS             <1. .>354
                     /codon_start=1
                     /product="antibody heavy chain variable region"
                     /db_xref="PID:e264666"
                     /db_xref="PID:g1518295"
                     /translation="EVKLQQSGPELVKPGASVKMSCKASGYTFTDYYMKWVKQSHGKS
                     LEWIGDINPNNGDTFYNQKFKGKATLTVDKSSSTAYMQLNSLTSEDSAVYYCANDGYH
                     WYFDVWGAGTPVTVSS"
BASE COUNT       92 a     89 c     93 g     80 t
ORIGIN

  Query Match           75.3%;  Score 262;  DB 28;  Length 354;
  Best Local Similarity  89.1%;  Pred. No. 3.24e-186;
  Matches   310;  Conservative    0;  Mismatches  36;  Indels   2;  Gaps   2;

Db       4 GTGAAGCTGCAGCAGTCGGGACCTGAGCTGGTGAAGCCTGGGGCTTCAGTGAAGATGTCC 63
           ||| |||||||| |||| |||||||| |||||||||||||| | |||| ||||||| |||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db      64 TGCAAGGCTTCTGGATACACCTTCACTGACTACTACATGAAGTGGGTGAAGCAGAGCCAT 123
           |||||| ||||||| ||| | ||||||| | ||  |||||| ||||||||||||||||||
Qy     425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db     124 GGAAAGAGCCTTGAGTGGATTGGAGATATTAATCCTAACAATGGTGATACTTTCTACAAC 183
           ||||||| ||||||||||||||||  |||||||||| ||||||||||||||  |||||||
Qy     485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db     184 CAGAAGTTCAAGGGCAAGGCCACATTGACTGTAGACAAATCCTCCAGCACAGCCTACATG 243
           |||||||||||||||||||||||||| |||||||||| || ||||||||||||||||||
Qy     545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db     244 CAGCTCAACAGCCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCCAATGATGGTTAC 303
            |||||  ||| ||||||||||||||||||||||||||||||||||| || || ||||||
Qy     605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGC-AAGGAGGGTTAC 663

Db     304 C-ACTGGTACTTCGATGTCTGGGGCGCAGGCACCCCCGTCACCGTCTC 350
             |||||||||||||||||||||||||||| ||| | ||||||||||
Qy     664 GGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT    7
LOCUS       S74055        355 bp    mRNA            ROD       01-MAR-1995
DEFINITION  IgG VH IIA-IgG2b VH region [mice, transfected Sp2/0 myeloma cells,
            mRNA Recombinant Partial, 355 nt].
ACCESSION   S74055
NID         g688162
KEYWORDS    .
SOURCE      Mus sp. transfected Sp2/0 myeloma cells.
  ORGANISM  Mus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 355)
  AUTHORS   Mount,P.F., Sutton,V.R., Li,W., Burgess,J., McKEnzie,I.F.,
            Pietersz,G.A. and Trapani,J.A.
  TITLE     Chimeric (mouse/human) anti-colon cancer antibody c30.6 inhibits
            the growth of human colorectal cancer xenografts in scid/scid mice
  JOURNAL   Cancer Res. 54 (23), 6160-6166 (1994)
  MEDLINE   95042371
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 157470] from the original journal article.
            This sequence comes from Fig. 1A.
FEATURES             Location/Qualifiers
     source          1. .355
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
     gene            1. .354
                     /partial
                     /note="IgG2b VH region"
                     /gene="IgG V<down>H</down> IIA"
     CDS             1. .354
                     /partial
                     /gene="IgG VH IIA"
                     /note="chimeric anti-colon cancer monoclonal antibody
                     c30.6/IgG heavy chain variable region; chimeric
                     mouse/human_.  This sequence comes from Fig. 1A"
                     /codon_start=1
                     /product="IgG2b VH region"
                     /db_xref="PID:g688163"
                     /translation="QVQLQESGPELVKPGASVKISCKASGYTFTDYNMHWVKQSHGKS
                     LEWIGFIYPYNAGTGYNQKFKNKATLTVDSSSSTAYMELRSLTSEDSAVYYCTRNDPH
                     WYFVVWGQGTTVTVSS"
BASE COUNT       93 a     92 c     90 g     80 t
ORIGIN

  Query Match           74.4%;  Score 259;  DB 28;  Length 355;
  Best Local Similarity  87.3%;  Pred. No. 1.00e-183;
  Matches   303;  Conservative    0;  Mismatches  44;  Indels   0;  Gaps   0;

Db       4 GTCCAACTGCAGGAGTCAGGGCCTGAGCTGGTGAAACCTGGGGCCTCAGTGAAGATATCC 63
           || || |||||||||||||| |||||| ||||||| ||||| |  ||| ||||||||||||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db      64 TGCAAGGCTTCTGGATACACATTCACTGACTACAACATGCACTGGGTGAAGCAGAGCCAT 123
           |||||| ||||||| ||| |||||||||| | ||| |||| ||||||||||||||||||||
Qy     425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db     124 GGAAAGAGTCTTGAGTGGATTGGATTTATTTATCCTTACAATGCTGGTACTGGCTACAAC 183
           |||||||  |||||||||||||||| ||||| |||||||||||| || ||||  |||||||
Qy     485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db     184 CAGAAATTCAAGAACAAGGCCACATTGACTGTAGACAGTTCCTCCAGCACAGCCTACATG 243
           ||||| ||||||  |||||||||||| ||||||||||  || ||||||||||||||||||
Qy     545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db     244 GAGCTCCGCAGCCTGACATCTGAGGACTCTGCAGTCTATTACTGTACAAGGAATGACCCC 303
           ||||||| ||| |||||||||||||||||||||||||||||||||| ||||||  |   |
Qy     605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTACG 664

Db     304 CACTGGTACTTCGTTGTCTGGGGCCAAGGGACCACGGTCACCGTCTC 350
           |||||||||||| ||||||||||  ||||||||||||||||||||||
Qy     665 GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT    8
LOCUS       MUSCMVG       357 bp    mRNA            ROD       30-JUN-1993
DEFINITION  Mouse mRNA sequence, partial cds.
```

```
ACCESSION   M95944
NID         g309176
KEYWORDS    .
SOURCE      Mus musculus (strain BALB/c, sub_species domesticus) cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 357)
  AUTHORS   Simpson,J.A., Chow,J.C., Baker,J., Avdalovic,N.M., Yuan,S.,
            Co,M.S., Vasquez,M., Britt,W.J. and Coelingh,K.L.
  TITLE     Untitled
  JOURNAL   Unpublished (1992)
FEATURES             Location/Qualifiers
     source          1. .357
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /sub_species="domesticus"
                     /db_xref="taxon:10090"
                     /cell_line="MAb 5"
                     /cell_type="hybridoma"
     mat_peptide     1. .357
                     /note="putative"
BASE COUNT       96 a     87 c     88 g     86 t
ORIGIN

  Query Match           74.4%;  Score 259;  DB 28;  Length 357;
  Best Local Similarity  89.1%;  Pred. No. 1.00e-183;
  Matches   312;  Conservative    0;  Mismatches  35;  Indels   3;  Gaps   2;

Db      4 GTCCAGCTGCAACAGTCTGGACCTGAGCTGGTGAAGCCTGGAGCTTCAATGAAGATATCC 63
          || ||||||||  |||| ||||||||| |||||||||||||||| ||||||||||||||||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     64 TGCAAGGCTTCTGTTTACTCATTCACTGGCTACACCATGAACTGGGTGAAGCAGAGCCAT 123
          |||||| |||||| |||||||||||||||| ||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    124 GGACAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGGTACTAGCTACAAC 183
          ||| ||||||||||||||||||||||||||||||||||||||||| ||||| |||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    184 CAGAAGTTCAAGGGGAAGGCCACATTAACTGTAGACAAGTCATCCAACACAGCCTACATG 243
          |||||||||||||| |||||||||| |||||||||||||| |||| ||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    244 GAGCTCCTCAGTCTGACATCTGCGGACTCTGCAGTCTATTACTGTACAAGACGGGGGTTT 303
          |||||||||||||||||||| |||||||||||||||||||||| ||||  | |||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG--GAGGGTTA 662

Db    304 CGAGACTATTCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTC 353
          || ||||  |   | ||   ||||||   ||| ||| | ||||||||||
Qy    663 CG-GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT    9
LOCUS       I78572        414 bp    DNA             PAT       23-DEC-1997
DEFINITION  Sequence 84 from patent US 5693761.
ACCESSION   I78572
NID         g3014726
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 414)
  AUTHORS   Queen,C.L., Schneider,W.P. and Selick,H.E.
  TITLE     Polynucleotides encoding improved humanized immunoglobulins
  JOURNAL   Patent: US 5693761-A 84 02-DEC-1997;
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="unknown"
BASE COUNT      105 a    102 c    103 g    104 t
ORIGIN

  Query Match           74.4%;  Score 259;  DB 21;  Length 414;
  Best Local Similarity  89.1%;  Pred. No. 1.00e-183;
  Matches   312;  Conservative    0;  Mismatches  35;  Indels   3;  Gaps   2;

Db     61 GTCCAGCTGCAACAGTCTGGACCTGAGCTGGTGAAGCCTGGAGCTTCAATGAAGATATCC 120
          || ||||||||  |||| ||||||||| |||||||||||||||| ||||||||||||||||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db    121 TGCAAGGCTTCTGTTTACTCATTCACTGGCTACACCATGAACTGGGTGAAGCAGAGCCAT 180
          |||||| |||||| |||||||||||||||| ||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    181 GGACAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGGTACTAGCTACAAC 240
          ||| ||||||||||||||||||||||||||||||||||||||||| ||||| |||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    241 CAGAAGTTCAAGGGGAAGGCCACATTAACTGTAGACAAGTCATCCAACACAGCCTACATG 300
          |||||||||||||| |||||||||| |||||||||||||| |||| ||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    301 GAGCTCCTCAGTCTGACATCTGCGGACTCTGCAGTCTATTACTGTACAAGACGGGGGTTT 360
          |||||||||||||||||||| |||||||||||||||||||||| ||||  | |||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG--GAGGGTTA 662

Db    361 CGAGACTATTCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTC 410
          || ||||  |   | ||   ||||||   ||| ||| | ||||||||||
Qy    663 CG-GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT   10
LOCUS       I78627        414 bp    DNA             PAT       23-DEC-1997
DEFINITION  Sequence 84 from patent US 5693762.
ACCESSION   I78627
NID         g3014781
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 414)
  AUTHORS   Queen,C.L., Co,M.Sung, Schneider,W.P., Landolfi,N.F., Coelingh,K.L.
            and Selick,H.E.
  TITLE     Humanized immunoglobulins
  JOURNAL   Patent: US 5693762-A 84 02-DEC-1997;
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="unknown"
BASE COUNT      105 a    102 c    103 g    104 t
ORIGIN

  Query Match           74.4%;  Score 259;  DB 21;  Length 414;
  Best Local Similarity  89.1%;  Pred. No. 1.00e-183;
  Matches   312;  Conservative    0;  Mismatches  35;  Indels   3;  Gaps   2;

Db     61 GTCCAGCTGCAACAGTCTGGACCTGAGCTGGTGAAGCCTGGAGCTTCAATGAAGATATCC 120
          || ||||||||  |||| ||||||||| |||||||||||||||| ||||||||||||||||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db    121 TGCAAGGCTTCTGTTTACTCATTCACTGGCTACACCATGAACTGGGTGAAGCAGAGCCAT 180
          |||||| |||||| |||||||||||||||| ||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    181 GGACAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGGTACTAGCTACAAC 240
          ||| ||||||||||||||||||||||||||||||||||||||||| ||||| |||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    241 CAGAAGTTCAAGGGGAAGGCCACATTAACTGTAGACAAGTCATCCAACACAGCCTACATG 300
          |||||||||||||| |||||||||| |||||||||||||| |||| ||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    301 GAGCTCCTCAGTCTGACATCTGCGGACTCTGCAGTCTATTACTGTACAAGACGGGGGTTT 360
          |||||||||||||||||||| |||||||||||||||||||||| ||||  | |||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG--GAGGGTTA 662
```

```
Db    361 CGAGACTATTCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTC 410
          || ||||  |   | ||   ||||||   ||| ||| | ||||||||||
Qy    663 CG-GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT   11
LOCUS       I31960        414 bp    DNA             PAT       07-JAN-1997
DEFINITION  Sequence 84 from patent US 5585089.
ACCESSION   I31960
NID         g1822751
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 414)
  AUTHORS   Queen,C.L. and Selick,H.E.
  TITLE     Humanized immunoglobulins
  JOURNAL   Patent: US 5585089-A 84 17-DEC-1996;
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="unknown"
BASE COUNT      105 a    102 c    103 g    104 t
ORIGIN

  Query Match           74.4%;  Score 259;  DB 21;  Length 414;
  Best Local Similarity  89.1%;  Pred. No. 1.00e-183;
  Matches   312;  Conservative    0;  Mismatches  35;  Indels   3;  Gaps   2;

Db     61 GTCCAGCTGCAACAGTCTGGACCTGAGCTGGTGAAGCCTGGAGCTTCAATGAAGATATCC 120
          || ||||||||  |||| |||||||||| |||||||||||||| ||||||||||||||||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db    121 TGCAAGGCTTCTGTTTACTCATTCACTGGCTACACCATGAACTGGGTGAAGCAGAGCCAT 180
          |||||| |||||| |||||||||||||| ||||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    181 GGACAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGGTACTAGCTACAAC 240
          ||| ||||||||||||||||||||||||||||||||||||||||||| ||||| |||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    241 CAGAAGTTCAAGGGGAAGGCCACATTAACTGTAGACAAGTCATCCAACACAGCCTACATG 300
          |||||||||||||| |||||||||| |||||||||||||| |||| ||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    301 GAGCTCCTCAGTCTGACATCTGCGGACTCTGCAGTCTATTACTGTACAAGACGGGGGTTT 360
          |||||||||||||||||||| |||||||||||||||||||||||| ||||  | |||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG--GAGGGTTA 662

Db    361 CGAGACTATTCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTC 410
          || ||||  |   | ||   ||||||   ||| ||| | ||||||||||
Qy    663 CG-GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT   12
LOCUS       S51851        421 bp                    ROD       08-MAR-1993
DEFINITION  M104E myeloma immunoglobulin heavy chain [mice, Genomic/mRNA, 421
            nt].
ACCESSION   S51851
NID         g262657
KEYWORDS    .
SOURCE      Mus sp.
  ORGANISM  Mus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 421)
  AUTHORS   Takahashi,S., Matsuura,Y., Taniguchi,T., Tamura,H., Bitoh,S.,
            Onishi,S., Yamamoto,Y., Yamamoto,H. and Fujimoto,S.
  TITLE     Molecular analysis of immunoglobulin heavy chain genes coding for
            idiotypic and anti-idiotypic antibodies involved in B-B cellular
            interaction
  JOURNAL   Microbiol. Immunol. 36 (8), 855-863 (1992)
  MEDLINE   93116638
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 121470] from the original journal article.
            This sequence comes from Fig. 1.
FEATURES             Location/Qualifiers
     source          1. .421
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
     gene            1. .421
                     /partial
                     /note="the leader peptide, VH region, D segment, and J
                     segment are shown completely; the constant region is shown
                     partially"
                     /gene="M104E myeloma immunoglobulin heavy chain"
     CDS             1. .421
                     /partial
                     /gene="M104E myeloma immunoglobulin heavy chain"
                     /codon_start=1
                     /translation="MGWSWIFLFLLSGTAGVLSEVQLQQSGPELVKPGASVKMSCKAS
                     GYTFTDYYMKWVKQSHGKSLEWIGDINPNNGGTSYNQKFKGKATLTVDKSSSTAYMQL
                     NSLTSEDSAVYYCARDYDWYFDVWGAGTTVTVSSESQS"
BASE COUNT      106 a    100 c    113 g    102 t
ORIGIN

  Query Match           74.1%;  Score 258;  DB 28;  Length 421;
  Best Local Similarity  89.3%;  Pred. No. 6.75e-183;
  Matches   310;  Conservative    0;  Mismatches  34;  Indels   3;  Gaps   2;

Db     61 GTCCAGCTGCAACAATCTGGACCTGAGCTGGTGAAGCCTGGGGCTTCAGTGAAGATGTCC 120
          || ||||||||  | || |||||||| |||||||||||||| | |||| ||||||| |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db    121 TGTAAGGCTTCTGGATACACATTCACTGACTACTACATGAAGTGGGTGAAGCAGAGTCAT 180
          || ||| ||||||| ||| |||||||||  | ||  |||||| |||||||||||||| |||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    181 GGAAAGAGCCTTGAGTGGATTGGAGATATTAATCCTAACAATGGTGGTACTAGCTACAAC 240
          ||||||| |||||||||||||||  |||||||||| |||||||| ||||| |||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    241 CAGAAGTTCAAGGGCAAGGCCACATTGACTGTAGACAAATCCTCCAGCACAGCCTACATG 300
          |||||||||||||||||||||||||| |||||||||| || ||||||||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    301 CAGCTCAACAGCCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGG--GATTACG 358
          |||||  ||| ||||||||||||||||||||||||||||||||||||||  | |||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTACG 664

Db    359 -ACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 404
           ||||||||||||||||||||||||||||||||||||||||||||||
Qy    665 GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT   13
LOCUS       AF035617      723 bp    mRNA            ROD       02-MAR-1998
DEFINITION  Mus musculus scFv antibody SS mRNA, partial cds.
ACCESSION   AF035617
NID         g2921393
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 723)
  AUTHORS   Chowdhury,P.S.
  TITLE     SS scFv
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 723)
  AUTHORS   Chowdhury,P.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-NOV-1997) LMB, NCI, 37 Convent Drive, Bld. 37 Rm.
            4B20, Bethesda, MD 20892-4255, USA
FEATURES             Location/Qualifiers
```

```
     source          1. .723
                     /organism="Mus musculus"
                     /strain="Balb/c"
                     /db_xref="taxon:10090"
                     /tissue_type="spleen"
                     /clone_lib="phage display library made from spleen mRNA of
                     mice immunized with DNA"
     CDS             1. .>723
                     /codon_start=1
                     /product="scFv antibody SS"
                     /db_xref="PID:g2921394"
                     /translation="MQVQLQQSGPELEKPGASVKISCKASGYSFTGYTMNWVKQSHGK
                     SLEWIGLITPYNGASSYNQKFRGKATLTVDKSSSTAYMDLLSLTSEDSAVYFCARGGY
                     DGRGFDYWGQGTTVTVSSGVGGSGGGGSGGGGSDIELTQSPAIMSASPGEKVTMTCSA
                     SSSVSYMHWYQQKSGTSPKRWIYDTSKLASGVPGRFSGSGSGNSYSLTISSVEAEDDA
                     TYYCQQWSGYPLTFGAGTKLEIK"
BASE COUNT      176 a    180 c    203 g    164 t
ORIGIN

  Query Match           73.9%;  Score 257;  DB 28;  Length 723;
  Best Local Similarity  88.9%;  Pred. No. 4.56e-182;
  Matches   311;  Conservative    0;  Mismatches  36;  Indels   3;  Gaps   2;

Db       7 GTACAACTGCAGCAGTCTGGGCCTGAGCTGGAGAAGCCTGGCGCTTCAGTGAAGATATCC 66
           || || |||||| |||| || |||||| ||| ||||||||| | |||| |||||||||||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db      67 TGCAAGGCTTCTGGTTACTCATTCACTGGCTACACCATGAACTGGGTGAAGCAGAGCCAT 126
           |||||| |||||||||||||||||||| |||||||||||||||||||||||||||||||||
Qy     425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db     127 GGAAAGAGCCTTGAGTGGATTGGACTTATTACTCCTTACAATGGTGCTTCTAGCTACAAC 186
           ||||||| |||||||||||||||||||||| |||||||||||||| | ||| |||||||
Qy     485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db     187 CAGAAGTTCAGGGGCAAGGCCACATTAACTGTAGACAAGTCATCCAGCACAGCCTACATG 246
           |||||||||| |||||||||||||| |||||||||||||| ||||||||||||||||||
Qy     545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db     247 GACCTCCTCAGTCTGACATCTGAAGACTCTGCAGTCTATTTCTGTGCAAGGGGGGGTTAC 306
           || |||||||||||||||||| |||||||||||||| |||||||||| ||| ||||
Qy     605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGG-TTAC 663

Db     307 GACGGGAGGGGTTTTGACTACTGGGGCCAAGGGACCACGGTCACCGTCTC 356
           | |   ||   || ||   |||||||  ||||||||||||||||||||
Qy     664 G--GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT   14
LOCUS       AF081541      366 bp    mRNA            ROD       31-AUG-1998
DEFINITION  Mus musculus anti-HIV-1 p24 single chain antibody A2 Ig heavy chain
            variable region mRNA, partial cds.
ACCESSION   AF081541
NID         g3493268
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 366)
  AUTHORS   de Haard,J.J.W., Kazemier,B., Koolen,M.J.M., Nijholt,L.J.,
            Meloen,R.H., van Gemen,B., Hoogenboom,H.R.M. and Arends,J.W.
  TITLE     Selection of recombinant, library derived antibody fragments
            against p24 for application in HIV-1 diagnostics
  JOURNAL   Clin. Diagn. Lab. Immunol. (1998) In press
REFERENCE   2  (bases 1 to 366)
  AUTHORS   de Haard,J.J.W., Kazemier,B., Koolen,M.J.M., Nijholt,L.J.,
            Meloen,R.H., van Gemen,B., Hoogenboom,H.R.M. and Arends,J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-1998) Functional Biomolecules, Unilever Research
            Laboratorium Vlaardingen, Olivier van Noortlaan 120, Vlaardingen
            3133 AT, The Netherlands
FEATURES              Location/Qualifiers
     source           1. .366
                      /organism="Mus musculus"
                      /strain="BALB/c"
                      /db_xref="taxon:10090"
                      /note="immunized with HIV-1 viral lysate"
     CDS              <1. .>366
                      /codon_start=1
                      /product="anti-HIV-1 p24 single chain antibody A2 Ig heavy
                      chain variable region"
                      /db_xref="PID:g3493269"
                      /translation="QVQLQQSGPELVKPGASVKLSCKLSGYSFTGYTMNWVKQSHGKN
                      LEWIGLINPYNGDTRNNQKFKGKATLTVDKSSSTAYMELLSLTSEDSAVYYCARDPYY
                      YGSSWYFDVWGAGTPVTVSS"
BASE COUNT       91 a     91 c     95 g     89 t
ORIGIN

  Query Match           72.4%;  Score 252;  DB 28;  Length 366;
  Best Local Similarity  93.8%;  Pred. No. 6.36e-178;
  Matches   270;  Conservative    0;  Mismatches  18;  Indels   0;  Gaps   0;

Db       4 GTTCAGCTGCAGCAGTCTGGGCCTGAGCTGGTGAAGCCTGGGGCTTCAGTGAAGTTGTCC 63
           || ||||||||| |||| || |||||| |||||||||||| | |||| ||||| | |||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db      64 TGCAAGCTTTCTGGTTACTCATTCACTGGCTACACCATGAACTGGGTGAAGCAGAGCCAT 123
           ||||||  |||||||||||||||||||| |||||||||||||||||||||||||||||||
Qy     425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db     124 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAGAAACAAC 183
           |||||||||||||||||||||||||||||||||||||||||||||||||||    |||||
Qy     485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db     184 CAGAAGTTCAAGGGCAAGGCCACATTAACTGTAGACAAGTCATCCAGCACAGCCTACATG 243
           |||||||||||||||||||||||| |||||||||||||| ||||||||||||||||||
Qy     545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db     244 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCA 291
           ||||||||||||||||||||||||||||||||||||||||||||||||
Qy     605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCA 652


RESULT   15
LOCUS       MUSHA4        440 bp    mRNA            ROD       24-JUL-1991
DEFINITION  Mouse active H-chain VJ region, 5' cds.
ACCESSION   M74137
NID         g193744
KEYWORDS    J-region; V-region; immunoglobulin heavy chain.
SOURCE      Mus musculus cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   Rueff-Juy,D., Marche,P.N., Drapier,A.-M. and Cazenave,P.-A.
  TITLE     Junctional diversity of H and L chains allows the coexpression of
            two mutually exclusive idiotopes (IdI104 and IdI558)
  JOURNAL   J. Immunol. 146, 4024-4030 (1991)
  MEDLINE   91237115
FEATURES              Location/Qualifiers
     source           1. .440
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /cell_line="hybridoma"
     CDS              39. .440
                      /partial
                      /codon_start=1
                      /product="Ab 16.3 heavy chain variable and joining
                      regions"
                      /db_xref="PID:g193745"
                      /translation="MELDLSLPLVRNCRCLSEVQLQQSGPELVKPGASVKMSCKASGY
                      TFTDYYMKWVKQSPGKSLEWIGDISPNNGGTSYNQKFKGKATLTVDKSSSTAYMQLNS
                      LTSEDSAVYYCARDYDWYFDVWGTGTTVTVSS"
BASE COUNT      114 a    112 c    110 g    104 t
```

ORIGIN

Query Match 71.8%; Score 250; DB 28; Length 440;
Best Local Similarity 88.2%; Pred. No. 2.88e-176;
Matches 306; Conservative 0; Mismatches 38; Indels 3; Gaps 3;

```
Db      93 GTCCAGCTGCAACAATCTGGACCTGAGCTGGTAAAGCCTGGGGCTTCAGTGAAGATGTCC 152
           || ||||||||  | || ||||||||| |||| |||||||| | |||| ||||||| |||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     153 TGTAAGGCTTCTGGATACACATTCACTGACTACTACATGAAGTGGGTGAAGCAGAGCCCT 212
           || ||| ||||||| ||| ||||||||| | ||  |||||| |||||||||||||||| |
Qy     425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db     213 GGAAAGAGCCTTGAGTGGATTGGAGATATTAGTCCTAACAATGGTGGTACTAGCTACAAC 272
           ||||||| |||||||||||||||  ||||| |||| |||||||| ||||| |||||||
Qy     485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db     273 CAGAAGTTCAAGGGCAAGGCCACATTGACTGTAGACAAATCCTCCAGCACAGCCTACATG 332
           |||||||||||||||||||||||||| |||||||||| || ||||||||||||||||||
Qy     545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db     333 CAGCTCAACAGCCTTACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG-AGA-TTACG 390
           |||||  ||| || ||||||||||||||||||||||||||||||||||| ||  |||||
Qy     605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTACG 664

Db     391 -ACTGGTACTTCGATGTCTGGGGCACAGGGACCACGGTCACCGTCTC 436
            |||||||||||||||||||||| ||||||||||||||||||||||
Qy     665 GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711
```

Search completed: Sat Jan 9 15:24:33 1999
Job time : 540 secs.

```
**************************************************************************
 ___   ___   _____   ____   _____   ____   _   _
| |\ \ / /| |  __| |  _ \ |  ___| |  _ \ | | | |
| | \ V / | | |__  | |_| | | |__  | |_| | | | | |
| |  \_/  | |  __| |  _ / |  __|  |  _ /  | |_| |
| |       | | |___ | | \ \ | |___ | | \ \  |  _  |
|_|       |_| |____||_|  \_\|_____||_|  \_\ |_| |_|  (TM)
**************************************************************************

        Release 3.1A John F. Collins, Biocomputing Research Unit.
        Copyright (c) 1993-1998  University of Edinburgh, U.K.
              Distribution rights by Oxford Molecular Ltd

MPsrch_nn   n.a. -  n.a. database search, using Smith-Waterman algorithm

Run on:          Sat Jan  9 16:26:30 1999;  MasPar time 64.16 Seconds
                                            737.977 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-704-178-1
Description:     (364-711) from US08704178.seq (3 of 4)
Perfect Score:   348
N.A. Sequence:      364 TGTGCAGCTGCAGGAGTCAG.........GGGACCACGGTCACCGTCTC 711
     Comp:              ACACGTCGACGTCCTCAGTC.........CCCTGGTGCCAGTGGCAGAG

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        n-geneseq32
                   1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                   8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
                   14:part14 15:part15 16:part16 17:part17 18:part18
                   19:part19 20:part20 21:part21 22:part22 23:part23
                   24:part24 25:part25 26:part26 27:part27 28:part28
                   29:part29 30:part30 31:part31 32:part32 33:part33
                   34:part34 35:part35 36:part36 37:part37 38:part38
                   39:part39 40:part40

Statistics:      Mean 8.057;  Variance 4.765;  scale 1.691

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	348	100.0	711	29	T65006	Single-chain anti-erb	4.40e-225
2	344	98.9	711	18	T17728	Anti-erbB2 scFv cDNA.	3.29e-222
3	344	98.9	711	9	Q55180	Sequence encoding the	3.29e-222
4	265	76.1	462	22	T34542	Monoclonal anti-idiot	1.26e-165
5	265	76.1	462	21	T31541	3H1 heavy chain varia	1.26e-165
6	265	76.1	462	38	T99435	Anti-idiotype antibod	1.26e-165
7	257	73.9	417	7	Q43843	Chimeric 128.1 VH, mo	6.39e-160
8	257	73.9	10704	7	Q43846	Plasmid pAH4625.	6.39e-160
9	257	73.9	10844	7	Q43848	Plasmid pAH4808.	6.39e-160
10	257	73.9	11529	7	Q43844	Plasmid pAH4602.	6.39e-160
11	257	73.9	12132	7	Q43847	Plasmid pAH4807.	6.39e-160
12	249	71.6	443	11	Q45428	KM-750 heavy chain.	3.21e-154
13	245	70.4	356	5	Q28522	Hypercalcaemia agent	2.27e-151
14	242	69.5	717	37	T86309	Single chain anti-dis	3.10e-149
15	242	69.5	1175	37	T86311	Single chain anti-dis	3.10e-149
16	240	69.0	913	12	Q81076	Bispecific CD3-L6FvIg	8.21e-148
17	235	67.5	354	11	Q65553	Mouse anti-HIV mu5.5	2.95e-144
18	235	67.5	354	12	Q70371	Anti HIV antibody hea	2.95e-144
19	231	66.4	449	11	Q45426	KM-796 heavy chain.	2.06e-141
20	231	66.4	1773	38	T88869	H chain subunit of Fa	2.06e-141
21	230	66.1	729	12	Q73679	Fv(GP-4) immunosuppre	1.06e-140
22	225	64.7	331	30	T43738	Anti-DNA antibody 8d8	3.76e-137
23	225	64.7	334	30	T43737	Anti-DNA antibody 7b3	3.76e-137
24	225	64.7	360	12	Q69933	U7.6 heavy chain vari	3.76e-137
25	225	64.7	360	19	T12612	scFv U7.6 VH gene.	3.76e-137
26	225	64.7	405	5	Q30754	p12-h2.	3.76e-137
27	225	64.7	474	2	Q12014	Sequence encoding mou	3.76e-137
28	225	64.7	474	2	Q12058	Sequence encoding hea	3.76e-137
29	224	64.4	424	24	T42717	HuMc3 VH coding seque	1.93e-136
30	222	63.8	426	40	V20602	A77 anti-Fc alpha R a	5.07e-135
31	221	63.5	372	16	T05018	FB5 antibody heavy ch	2.60e-134
32	221	63.5	402	25	T31293	Guy's 13 anti-Strepto	2.60e-134
33	221	63.5	450	1	Q04695	Heavy chain variable	2.60e-134
34	220	63.2	420	40	V09796	DNA encoding the heav	1.33e-133
35	219	62.9	360	40	V09798	DNA encoding the heav	6.82e-133
36	218	62.6	413	24	T34153	Monoclonal antibody P	3.49e-132
37	217	62.4	357	18	T04014	Anti-EGFR antibody he	1.79e-131
38	217	62.4	387	4	Q27141	ICAM-1 inhibiting pep	1.79e-131
39	217	62.4	486	31	T59339	MH1 monoclonal antibo	1.79e-131
40	215	61.8	354	18	T04018	Anti-EGFR antibody he	4.68e-130
41	213	61.2	357	31	T58254	Lead binding MAb 14F1	1.22e-128
42	212	60.9	405	33	T78589	Monoclonal antibody 6	6.25e-128
43	212	60.9	499	2	Q11291	Encodes murine monocl	6.25e-128
44	212	60.9	762	17	T03381	Anti-IL-8 MAb 6G4.2.5	6.25e-128
45	212	60.9	762	33	T78595	Chimeric monoclonal a	6.25e-128

ALIGNMENTS

```
RESULT    1
ID   T65006 standard; cDNA; 711 BP.
AC   T65006;
DT   05-JUN-1997  (first entry)
DE   Single-chain anti-erbB2 antibody e23(Fv) cDNA.
KW   Single chain antibody; variable region; light chain; heavy chain;
KW   breast cancer; ovarian cancer; non-small cell lung carcinoma;
KW   immunodiagnosis; treatment; cytotoxic agent; erbB-2; ds.
OS   Mus musculus.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     1..711
FT                   /*tag=  a
FT                   /product= e23(Fv)
PN   US5587458-A.
PD   24-DEC-1996.
PF   07-OCT-1991; 772270.
PR   07-OCT-1991; US-772270.
PR   30-JUN-1992; US-906555.
PR   14-MAY-1993; US-061092.
PA   (ARON-) ARONEX PHARM INC.
PI   Bird RE,  Kasprzyk PG,  King CR;
DR   WPI; 97-064831/06.
DR   P-PSDB; W15185.
PT   Single chain antibodies specific for erbB-2 protein, gp185  - with
PT   labels or cytotoxin, useful for detection and treatment of tumour
PT   cells expressing this protein
PS   Example 8; Columns 25-28; 28pp; English.
CC   The present cDNA sequence codes for a claimed single-chain antibody,
CC   designated e23(Fv), which binds to erbB-2. Monoclonal antibody e23
CC   was generated by immunising mice with N/erbB-2 cells overexpressing
CC   the gp185 protein, removing spleen cells and producing hybridomas
CC   by standard techniques. Messenger RNA coding for the anti-erbB-2
CC   monoclonal antibody was isolated and converted to cDNA. Regions
CC   coding for the heavy- and light- chain variable regions were then
CC   amplified by PCR and joined via a sequence encoding a peptide
CC   linker. The resulting single-chain antibody is useful for in vitro
```

```
CC   diagnosis of tumour cells which overexpress the erbB-2 gp185
CC   marker, e.g. breast, ovarian and non-small cell lung carcinomas,
CC   and, when coupled to a cytotoxic agent, to treat such tumours.
SQ   Sequence   711 BP;    175 A;    182 C;    189 G;    164 T;

  Query Match          100.0%;  Score 348;  DB 29;  Length 711;
  Best Local Similarity 100.0%;  Pred. No. 4.40e-225;
  Matches   348;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    363 tgtgcagctgcaggagtcaggacctgaggtggtgaagcctggaggttcaatgaagatatc 422
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    364 TGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATC 423

Db    423 ctgcaagacttctggttactcattcactggccacaccatgaactgggtgaagcagagcca 482
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    424 CTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCA 483

Db    483 tggaaagaaccttgagtggattggacttattaatccttacaatggtgatactaactacaa 542
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    484 TGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAA 543

Db    543 ccagaagttcaagggcaaggccacatttactgtagacaagtcgtccagcacagcctacat 602
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    544 CCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACAT 603

Db    603 ggagctcctcagtctgacatctgaggactctgcagtctattactgtgcaaggagggttac 662
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    604 GGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTAC 663

Db    663 ggactggtacttcgatgtctggggcgcagggaccacggtcaccgtctc 710
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    664 GGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT    2
ID   T17728 standard; cDNA; 711 BP.
AC   T17728;
DT   21-MAY-1996  (first entry)
DE   Anti-erbB2 scFv cDNA.
KW   Oncoprotein; erbB2; cell proliferation; tumour; cancer;
KW   intracellular antibody homologue; single chain antibody; scFv;
KW   gene therapy; ds.
OS   Synthetic.
PN   WO9607321-A1.
PD   14-MAR-1996.
PF   23-AUG-1995; U10740.
PR   06-SEP-1994; US-301339.
PR   06-JUN-1995; US-468252.
PA   (UABR-) UAB RES FOUND.
PI   Curiel DT,  Deshane J;
DR   WPI; 96-171307/17.
DR   P-PSDB; R94020.
PT   Inhibition of proliferation or survival of, esp. malignant erbB2,
PT   cells - by introducing nucleic acid mol. encoding antibody homologue
PT   which is expressed and binds, pref. erbB2, protein intracellularly
PS   Claim 42; Page 29-30;. 48pp; English.
CC   A nucleic acid comprises a first sequence encoding a signal peptide
CC   (R94019) linked to a second sequence (T17728) encoding a single
CC   chain Fv fragment (R94020) that binds a human erbB2 oncoprotein.
CC   The anti-erbB2 sFv portion is obtained by PCR using e23scFv
CC   plasmid as template.  The signal peptide directs the scFv to the
CC   endoplasmic reticulum.  The nucleic acid is incorporated into a
CC   plasmid or viral vector to facilitate expression of the scFv antibody
CC   homologue within e.g. an epithelial carcinoma cell.  Intracellular
CC   expression of the homologue inhibits surface expression of erbB2 and
CC   thereby inhibits cell proliferation and cell survival and decreases
CC   tumorigenicity.
SQ   Sequence   711 BP;    175 A;    183 C;    189 G;    164 T;

  Query Match          98.9%;  Score 344;  DB 18;  Length 711;
  Best Local Similarity  99.4%;  Pred. No. 3.29e-222;
  Matches   346;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Db    363 tgtgcagctgcaggagtcaggacctgaggtggtcaagcctggaggttcaatgaagatatc 422
        |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy    364 TGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATC 423

Db    423 ctgcaagacttctggttactcattcactggccacaccatgaactgggtgaagcagagcca 482
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    424 CTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCA 483

Db    483 tggaaagaaccttgagtggattggacttattaatccttacaatggtgatactaactacaa 542
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    484 TGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAA 543

Db    543 ccagaagttcaagggcaaggccacatttactgtagacaagtcgtccagcacagcctacat 602
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    544 CCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACAT 603

Db    603 ggagctcctcagtctgacatctgaggactctgcagtctattactgtgcaaggagggttac 662
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    604 GGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTAC 663

Db    663 ggactggtacttcgatgtctgggggcagggaccacggtcaccgtctc 710
        ||||||||||||||||||||||| ||||||||||||||||||||||||
Qy    664 GGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT    3
ID   Q55180 standard; cDNA; 711 BP.
AC   Q55180;
DT   21-JUL-1994  (first entry)
DE   Sequence encoding the single chain anti-erbB2 antibody, Ab no.23.
KW   Single chain anti-erbB1 antibody; cancer therapy; prevention;
KW   monoclonal antibody; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   cds             1..171
FT                   /*tag=  a
PN   WO9400136-A.
PD   06-JAN-1994.
PF   21-OCT-1992; U08545.
PR   30-JUN-1992; US-906555.
PA   (MOLE-) MOLECULAR ONCOLOGY INC.
PI   Kasprzyk PG,  King CR;
DR   WPI; 94-025878/03.
DR   P-PSDB; R45442.
PT   Treatment of malignancies over-expressing ERB-[2 - using at least
PT   2 monoclonal antibodies which recognise different epitopes on
PT   gp185
PS   Example; Fig 7; 37pp; English.
CC   The source of human erbB-2 protein for the prodn. of antibodies no.
CC   23 and 21 (Q55180/R45442; Q55181/R45443) is a NIH/3T3 cell
CC   engineered to express the human erbB-2 protein on its surface (N/
CC   erbB-2). Abs no. 21 and 23 are directed against the extracellular
CC   domain of gp185 erbB-2. Nude mice manipulated to produce rapidly
CC   growing tumours were used in a trial of the efficacy of the Abs. In
CC   animals given a combination of the 2 Abs, tumours completely
CC   regressed after 11 days.
SQ   Sequence   711 BP;    175 A;    183 C;    187 G;    165 T;

  Query Match          98.9%;  Score 344;  DB 9;  Length 711;
  Best Local Similarity  99.4%;  Pred. No. 3.29e-222;
  Matches   346;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Db    363 tgtgcagctgcaggattcaggacctgaggtggtgaagcctggaggttcaatgaagatatc 422
        |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Qy    364 TGTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATC 423

Db    423 ctgcaagacttctggttactcattcactggccacaccatgaactgggtgaagcagagcca 482
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    424 CTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCA 483

Db    483 tggaaagaaccttgagtggattggacttattaatccttacaatggtgatactaactacaa 542
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    484 TGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAA 543
```

```
Db    543 ccagaagttcaagggcaaggccacatttactgtagacaagtcgtccagcacagcctacat 602
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    544 CCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACAT 603

Db    603 ggagctcctcagtctgacatctgaggactctgcactctattactgtgcaaggagggttac 662
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Qy    604 GGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTAC 663

Db    663 ggactggtacttcgatgtctggggcgcagggaccacggtcaccgtctc 710
          |||||||||||||||||||||||||||||||||||||||||||||||
Qy    664 GGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT    4
ID   T34542 standard; cDNA; 462 BP.
AC   T34542;
DT   11-OCT-1996  (first entry)
DE   Monoclonal anti-idiotype antibody 3H1 VH cDNA.
KW   Anti-idiotype antibody; monoclonal antibody; CEA;
KW   carcinoembryonic antigen; tumour-associated antigen; cancer;
KW   vaccine; immunotherapy; ss.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   cds             22..462
FT                   /*tag=  a
PN   WO9620219-A2.
PD   04-JUL-1996.
PF   28-DEC-1995; U17105.
PR   28-DEC-1994; US-365484.
PA   (KENT ) UNIV KENTUCKY.
PI   Chatterjee M,  Chatterjee SK,  Foon KA,  Kohler H;
DR   WPI; 96-321809/32.
DR   P-PSDB; R99687.
PT   Monoclonal anti-idiotype antibody 3H1 - elicits an immune response
PT   to carcinoembryonic antigen
PS   Example 2; Fig 2A; 102pp; English.
CC   A cDNA clone (T34542) codes for the heavy chain variable region
CC   (R99687) of monoclonal anti-idiotype antibody 3H1 (ATCC HB 12003).
CC   It was obtd. by PCR amplification (see also T34543-44) of cDNA
CC   derived from 3H1 hybridoma cells.  Anti-idiotype antibody 3H1
CC   elicits a specific immune response to a unique epitope of
CC   carcinoembryonic antigen (CEA) that is not present on other
CC   members of the CEA family or on normal adult tissues.  3H1
CC   can be used as a vaccine to elicit immune responses in patients
CC   with advanced CEA-associated disease or, when labeled, to enhance
CC   tumour detection in imaging.
SQ   Sequence   462 BP;    106 A;    124 C;    117 G;    115 T;

  Query Match           76.1%;  Score 265;  DB 22;  Length 462;
  Best Local Similarity  90.0%;  Pred. No. 1.26e-165;
  Matches   315;  Conservative    0;  Mismatches  32;  Indels   3;  Gaps   2;

Db     82 gtccagctgcaacagtctggacctgagctggtgaagcctggagcttcactgaagatttcc 141
          || ||||||||  |||| |||||||||| ||||||||||||||| |||| ||||||| |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db    142 tgcgaggcttctggttactcactcactgcctacaccatgaactgggtgaagcagagccat 201
          ||| || ||||||||||||||| |||||| | ||||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    202 ggaaagagccttgagtgggttgggctgattaatcctttcagtggtgatactaactacagc 261
          |||||||| |||||||||| |||| || |||||||||| || ||||||||||||||||| |
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    262 cagaaattcacgggcaaggccacattaactgtagacaggtcatccagcacagcctacatg 321
          ||||| |||| |||||||||||||| |||||||||| ||| ||||||||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    322 gagctcctcagtctgacatctgaggactctgcagtctattactgtgtcattactccggtt 381
          |||||||||||||||||||||||||||||||||||||||||||||||| ||  |    ||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTG-CA--AGGAGGGTT 661

Db    382 ccctactggtacttcgatgtctggggcgcagggaccacggtcaccgtctc 431
          |  ||||||||||||||||||||||||||||||||||||||||||||||
Qy    662 ACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT    5
ID   T31541 standard; cDNA; 462 BP.
AC   T31541;
DT   18-SEP-1996  (first entry)
DE   3H1 heavy chain variable region cDNA.
KW   Anti-idiotype; monoclonal antibody; MAb; 3H1; CEA;
KW   carcinoembryonic antigen; cancer; gene therapy; immunotherapy;
KW   vaccine; genetic immunisation; heavy chain; VH; ss.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   cds             22..462
FT                   /*tag=  a
FT                   /product= 3H1 heavy chain variable region
PN   WO9620277-A2.
PD   04-JUL-1996.
PF   28-DEC-1995; U17103.
PR   28-DEC-1994; US-365484.
PA   (KENT ) UNIV KENTUCKY.
PI   Chatterjee M,  Chatterjee SK,  Foon KA,  Kohler H;
DR   WPI; 96-321850/32.
DR   P-PSDB; R98411.
PT   Recombinant monoclonal anti-idiotype antibody 3H1 sequences - used
PT   to develop prods. for the detection and treatment of
PT   carcinoembryonic antigen-associated diseases, partic. cancers
PS   Claim 6; Fig 2A; 121pp; English.
CC   A cDNA clone (T31541) codes for the heavy chain variable region
CC   (VH) (R98411) of 3H1, a murine monoclonal anti-idiotype antibody
CC   that mimics a specific epitope of the 180,000 mol.wt.
CC   carcinoembryonic antigen (CEA) and which elicits an immune response
CC   in patients with advanced CEA-associated disease, e.g. colorectal
CC   cancer.  It was obtd. by PCR amplification of 3H1 hybridoma DNA
CC   (see also T31545-46).  The isolated polynucleotide, and/or a
CC   polynucleotide (see also T31540) coding for the VL region (R98410)
CC   of 3H1 can be used to design probes and primers, in expression
CC   systems, and in pharmaceutical applns., including vaccines, gene
CC   therapy and genetic immunisation, partic. against cancer.
SQ   Sequence   462 BP;    106 A;    124 C;    117 G;    115 T;

  Query Match           76.1%;  Score 265;  DB 21;  Length 462;
  Best Local Similarity  90.0%;  Pred. No. 1.26e-165;
  Matches   315;  Conservative    0;  Mismatches  32;  Indels   3;  Gaps   2;

Db     82 gtccagctgcaacagtctggacctgagctggtgaagcctggagcttcactgaagatttcc 141
          || ||||||||  |||| |||||||||| ||||||||||||||| |||| ||||||| |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db    142 tgcgaggcttctggttactcactcactgcctacaccatgaactgggtgaagcagagccat 201
          ||| || ||||||||||||||| |||||| | ||||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    202 ggaaagagccttgagtgggttgggctgattaatcctttcagtggtgatactaactacagc 261
          |||||||| |||||||||| |||| || |||||||||| || ||||||||||||||||| |
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    262 cagaaattcacgggcaaggccacattaactgtagacaggtcatccagcacagcctacatg 321
          ||||| |||| |||||||||||||| |||||||||| ||| ||||||||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    322 gagctcctcagtctgacatctgaggactctgcagtctattactgtgtcattactccggtt 381
          |||||||||||||||||||||||||||||||||||||||||||||||| ||  |    ||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTG-CA--AGGAGGGTT 661

Db    382 ccctactggtacttcgatgtctggggcgcagggaccacggtcaccgtctc 431
          |  ||||||||||||||||||||||||||||||||||||||||||||||
Qy    662 ACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT    6
```

```
ID   T99435 standard; cDNA; 462 BP.
AC   T99435;
DT   27-APR-1998  (first entry)
DE   Anti-idiotype antibody 3H1 heavy chain variable region cDNA.
KW   Anti-idiotype antibody 3H1; carcinoembryonic antigen; CEA;
KW   tumour-specific antigen; tumour; colorectal cancer; lung cancer;
KW   adenocarcinoma; therapy; ss.
OS   Mus musculus.
FH   Key             Location/Qualifiers
FT   CDS             22..462
FT                   /*tag=  a
PN   WO9738725-A1.
PD   23-OCT-1997.
PF   11-APR-1997; U05953.
PR   09-APR-1997; US-631085.
PR   12-APR-1996; US-631085.
PA   (KENT ) UNIV KENTUCKY.
PI   Chatterjee M, Chatterjee SK, Foon KA;
DR   WPI; 97-526218/48.
DR   P-PSDB; W26729.
PT   Use of anti-idiotype antibody 3H1 - for delaying development of
PT   carcinoembryonic antigen associated tumours
PS   Disclosure; Page 39; 66pp; English.
CC   This cDNA sequence includes a coding region for the heavy chain
CC   variable region (VH) (see W26729) of murine monoclonal anti-idiotype
CC   antibody 3H1.  A cDNA sequence (see T99434) for the VL region (see
CC   W26728) is also provided.  These clones can be utilised in the
CC   recombinant production of 3H1 polypeptides.  3H1 induces a specific
CC   immune response against a distinct and specific epitope of
CC   carcinoembryonic antigen (CEA), a tumour-associated antigen.  The
CC   invention provides methods of delaying development of CEA-associated
CC   tumours using 3H1, particularly in high-risk individuals.  The
CC   method can be used to delay the development of a variety of tumours,
CC   including colorectal, gastric and pancreatic tumours, as well as
CC   other adenocarcinomas such as those of the breast and lung, and
CC   biliary cancer, and gynaecological cancers.
SQ   Sequence   462 BP;    106 A;    124 C;    117 G;    115 T;

  Query Match           76.1%;  Score 265;  DB 38;  Length 462;
  Best Local Similarity  90.0%;  Pred. No. 1.26e-165;
  Matches   315;  Conservative    0;  Mismatches  32;  Indels   3;  Gaps   2;

Db     82 gtccagctgcaacagtctggacctgagctggtgaagcctggagcttcactgaagatttcc 141
          || ||||||||  |||| |||||||| ||||||||||||||| |||| ||||||| |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db    142 tgcgaggcttctggttactcactcactgcctacaccatgaactgggtgaagcagagccat 201
          ||| || |||||||||||||| |||||| | ||||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    202 ggaaagagccttgagtgggttgggctgattaatcctttcagtggtgatactaactacagc 261
          ||||||| |||||||||| |||| || |||||||||| || |||||||||||||||| |
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    262 cagaaattcacgggcaaggccacattaactgtagacaggtcatccagcacagcctacatg 321
          ||||| |||| |||||||||||||| |||||||||| ||| ||||||||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    322 gagctcctcagtctgacatctgaggactctgcagtctattactgtgtcattactccggtt 381
          ||||||||||||||||||||||||||||||||||||||||||||| ||  |    ||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTG-CA--AGGAGGGTT 661

Db    382 ccctactggtacttcgatgtctggggcgcagggaccacggtcaccgtctc 431
           |  ||||||||||||||||||||||||||||||||||||||||||||||
Qy    662 ACGGACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT    7
ID   Q43843 standard; DNA; 417 BP.
AC   Q43843;
DT   20-OCT-1993  (first entry)
DE   Chimeric 128.1 VH, mouse gamma subgroup IIB DNA.
KW   Polymerase chain reaction; primer; PCR; amplify; murine; heavy;
KW   light; chain; variable; constant; region; anti-human; transferrin;
KW   receptor; antibody; brain; capillary; endothelial cell; conjugate;
KW   neuropharmaceutical; diagnostic; agent; tumour; AIDS; stroke;
KW   epilepsy; Parkinsons disease; Alzheimers disease; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   signal_peptide  1..57
FT                   /*tag=  a
FT                   /note= "Leader sequence"
FT   mat_peptide     58..417
FT                   /*tag=  b
PN   WO9310819-A.
PD   10-JUN-1993.
PF   24-NOV-1992; U10206.
PR   26-NOV-1991; US-800458.
PA   (ALKE-) ALKERMES INC.
PI   Friden PM;
DR   WPI; 93-196742/24.
DR   P-PSDB; R38259.
PT   Antibody conjugates specific for transferrin receptor - used for
PT   diagnosis and treatment of cancer, AIDS and neurological
PT   disorders
PS   Table 6; Page 57; 151pp; English.
CC   The sequences given in Q43842-43 encode the light and heavy chains,
CC   respectively, of the chimeric antibody 128.1.  128.1 is an anti-human
CC   transferrin receptor antibody which binds to the transferrin receptor
CC   on brain capillary endothelial cells.  This antibody may be used in a
CC   conjugate in which it is linked to a neuropharmaceutical or diagnostic
CC   agent.  The conjugate may be used to treat or prevent neurological
CC   disorders eg. brain tumours, AIDS, stroke, epilepsy, Parkinsons and
CC   Alzheimers disease.  It may also be used for diagnostic methods.
SQ   Sequence   417 BP;    108 A;    104 C;    102 G;    103 T;

  Query Match           73.9%;  Score 257;  DB 7;  Length 417;
  Best Local Similarity  87.0%;  Pred. No. 6.39e-160;
  Matches   302;  Conservative    0;  Mismatches  45;  Indels   0;  Gaps   0;

Db     61 gtccagctgcaacagtctggacctgaactggtgaagcctggagcttcaatgaagatttcc 120
          || ||||||||  |||| |||||||| |||||||||||||| |||||||||||| |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db    121 tgcaaggcttctggttactcattcactggctacaccatgaactgggtgaagcagagccat 180
          |||||| |||||||||||||||||||||| ||||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    181 ggagagaaccttgagtggattggacgtattaatcctcacaatggtggtactgactacaac 240
          ||| |||||||||||||||||||| |||||||||| ||||||||| |||| ||||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    241 cagaagttcaaggacaaggccccctttaactgtagacaagtcatccaacacagcctacatg 300
          |||||||||||| ||||||| | || |||||||||||||| |||| ||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    301 gagctcctcagtctgacatctgaggactctgcagtctattactgtgcaagaggctactat 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTACG 664

Db    361 tactattctttggactactggggtcaaggaacctcagtcaccgtctc 407
           |||  |  || ||   |||||||   ||| ||| | ||||||||||
Qy    665 GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT    8
ID   Q43846 standard; DNA; 10704 BP.
AC   Q43846;
DT   20-OCT-1993  (first entry)
DE   Plasmid pAH4625.
KW   Polymerase chain reaction; primer; PCR; amplify; murine; pAH4625;
KW   heavy; light; chain; variable; constant; region; anti-human; pAH4807;
KW   transferrin; receptor; antibody; brain; capillary; pAH4808; gamma-2;
KW   endothelial cell; conjugate; neuropharmaceutical; gamma-3; gamma-4;
KW   diagnostic; agent; tumour; AIDS; stroke; epilepsy; monoclonal;
KW   Parkinsons disease; Alzheimers disease; SP2/0 cell; ss.
```

```
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   unsure          11..1710
FT                   /*tag=  a
FT                   /note= "Sequence not given in the specification"
FT   cds             1848..2552
FT                   /*tag=  b
FT                   /note= "Chimeric 128.1 VH/CH1 region"
FT   cds             2945..2980
FT                   /*tag=  c
FT                   /note= "Undefined ORF1"
FT   cds             3099..3425
FT                   /*tag=  d
FT                   /note= "Undefined ORF2"
FT   unsure          4012..4311
FT                   /*tag=  e
FT                   /note= "Sequence not given in the specification"
FT   unsure          4318..4517
FT                   /*tag=  f
FT                   /note= "Sequence not given in the specification"
PN   WO9310819-A.
PD   10-JUN-1993.
PF   24-NOV-1992; U10206.
PR   26-NOV-1991; US-800458.
PA   (ALKE-) ALKERMES INC.
PI   Friden PM;
DR   WPI; 93-196742/24.
DR   P-PSDB; R41707-09.
PT   Antibody conjugates specific for transferrin receptor - used
PT   for diagnosis and treatment of cancer, AIDS and neurological
PT   disorders
PS   Disclosure; Fig 17A-F; 151pp; English.
CC   The sequences given in Q43846-48 represent the expression vectors
CC   pAH4625, pAH4807 and pAH4808.  These vectors represent the cloning
CC   of the different human gamma isotypes, gamma-2, gamma-3 and gamma-4
CC   respectively, with the variable region of the murine monoclonal
CC   antibody 128.1.  These plasmids each encode a chimeric monoclonal
CC   antibody in which the heavy chain (VH) is derived from a murine
CC   source and the sequences encoding CH1, CH2 and CH3 are derived from
CC   a human source.  These vectors in combination with the chimeric light
CC   chain vector, pAG4611 (see also Q43845), were transfected into SP2/0
CC   cells and clones were isolated.  Antibody analysis using biosynthe-
CC   tically labelled proteins, immunoprecipitation and SDS-PAGE indicated
CC   appropriate bands for the heavy and light chains as well as the
CC   assembled antibody for the gamma-3 and gamma-4 chimeras.  No detectable
CC   protein was isolated for the gamma-2 transfectants.  128.1 is an anti-
CC   human transferrin receptor antibody which binds to the transferrin
CC   receptor on brain capillary endothelial cells.  This antibody may be
CC   used in a conjugate in which it is linked to a neuropharmaceutical or
CC   diagnostic agent.  The conjugate may be used to treat or prevent
CC   neurological disorders eg. brain tumours, AIDS, stroke, epilepsy,
CC   Parkinsons and Alzheimers disease.  It may also be used for diagnostic
CC   methods.
SQ   Sequence   10704 BP;    2125 A;    2189 C;    2130 G;    2031 T;

  Query Match           73.9%;  Score 257;  DB 7;  Length 10704;
  Best Local Similarity  87.0%;  Pred. No. 6.39e-160;
  Matches   302;  Conservative    0;  Mismatches  45;  Indels   0;  Gaps   0;

Db   1908 gtccagctgcaacagtctggacctgaactggtgaagcctggagcttcaatgaagatttcc 1967
          || ||||||||  |||| ||||||||  ||||||||||||||| |||||||||||| |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db   1968 tgcaaggcttctggttactcattcactggctacaccatgaactgggtgaagcagagccat 2027
          |||||| ||||||||||||||||||||||| |||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db   2028 ggagagaaccttgagtggattggacgtattaatcctcacaatggtggtactgactacaac 2087
          ||| ||||||||||||||||||| |||||||||| ||||||||| |||| ||||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db   2088 cagaagttcaaggacaaggccccctttaactgtagacaagtcatccaacacagcctacatg 2147
          ||||||||||||| ||||||| | || |||||||||||||| |||| ||||||||||||

Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db   2148 gagctcctcagtctgacatctgaggactctgcagtctattactgtgcaagaggctactat 2207
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTACG 664

Db   2208 tactattctttggactactggggtcaaggaacctcagtcaccgtctc 2254
          |||  |  || ||   ||||||   ||| ||| | ||||||||||
Qy    665 GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT    9
ID   Q43848 standard; DNA; 10844 BP.
AC   Q43848;
DT   20-OCT-1993  (first entry)
DE   Plasmid pAH4808.
KW   Polymerase chain reaction; primer; PCR; amplify; murine; pAH4625;
KW   heavy; light; chain; variable; constant; region; anti-human; pAH4807;
KW   transferrin; receptor; antibody; brain; capillary; pAH4808; gamma-2;
KW   endothelial cell; conjugate; neuropharmaceutical; gamma-3; gamma-4;
KW   diagnostic; agent; tumour; AIDS; stroke; epilepsy; monoclonal;
KW   Parkinsons disease; Alzheimers disease; SP2/0 cell; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   unsure          2070..3769
FT                   /*tag=  a
FT   cds             3907..4611
FT                   /*tag=  b
FT                   /note= "Murine 128.1 VH/human gamma-4 CH1"
FT   cds             4999..5034
FT                   /*tag=  c
FT                   /note= "Undefined ORF1"
FT   cds             5153..5482
FT                   /*tag=  d
FT                   /note= "Undefined ORF2"
FT   cds             5580..5900
FT                   /*tag=  e
FT                   /note= "Undefined ORF3"
FT   unsure          6128..6329
FT                   /*tag=  f
FT   unsure          6336..6635
FT                   /*tag=  g
PN   WO9310819-A.
PD   10-JUN-1993.
PF   24-NOV-1992; U10206.
PR   26-NOV-1991; US-800458.
PA   (ALKE-) ALKERMES INC.
PI   Friden PM;
DR   WPI; 93-196742/24.
DR   P-PSDB; R41715-18.
PT   Antibody conjugates specific for transferrin receptor - used
PT   for diagnosis and treatment of cancer, AIDS and neurological
PT   disorders
PS   Disclosure; Fig 19A-F; 151pp; English.
CC   The sequences given in Q43846-48 represent the expression vectors
CC   pAH4625, pAH4807 and pAH4808.  These vectors represent the cloning
CC   of the different human gamma isotypes, gamma-2, gamma-3 and gamma-4
CC   respectively, with the variable region of the murine monoclonal
CC   antibody 128.1.  These plasmids each encode a chimeric monoclonal
CC   antibody in which the heavy chain (VH) is derived from a murine
CC   source and the sequences encoding CH1, CH2 and CH3 are derived from
CC   a human source.  These vectors in combination with the chimeric light
CC   chain vector, pAG4611 (see also Q43845), were transfected into SP2/0
CC   cells and clones were isolated.  Antibody analysis using biosynthe-
CC   tically labelled proteins, immunoprecipitation and SDS-PAGE indicated
CC   appropriate bands for the heavy and light chains as well as the
CC   assembled antibody for the gamma-3 and gamma-4 chimeras.  No detectable
CC   protein was isolated for the gamma-2 transfectants.  128.1 is an anti-
CC   human transferrin receptor antibody which binds to the transferrin
CC   receptor on brain capillary endothelial cells.  This antibody may be
CC   used in a conjugate in which it is linked to a neuropharmaceutical or
CC   diagnostic agent.  The conjugate may be used to treat or prevent
CC   neurological disorders eg. brain tumours, AIDS, stroke, epilepsy,
```

```
CC   Parkinsons and Alzheimers disease.  It may also be used for diagnostic
CC   methods.
SQ   Sequence   10844 BP;    2173 A;    2228 C;    2150 G;    2063 T;

  Query Match           73.9%;  Score 257;  DB 7;  Length 10844;
  Best Local Similarity  87.0%;  Pred. No. 6.39e-160;
  Matches   302;  Conservative    0;  Mismatches  45;  Indels   0;  Gaps   0;

Db   3967 gtccagctgcaacagtctggacctgaactggtgaagcctggagcttcaatgaagatttcc 4026
            || ||||||||  |||| ||||||||  ||||||||||||||| ||||||||||| |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db   4027 tgcaaggcttctggttactcattcactggctacaccatgaactgggtgaagcagagccat 4086
            |||||| |||||||||||||||||||||| |||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db   4087 ggagagaaccttgagtggattggacgtattaatcctcacaatggtggtactgactacaac 4146
            ||| |||||||||||||||||||| |||||||||| ||||||||| |||| ||||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db   4147 cagaagttcaaggacaaggccccttttaactgtagacaagtcatccaacacagcctacatg 4206
            |||||||||||||| ||||||| | || |||||||||||||| |||| ||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db   4207 gagctcctcagtctgacatctgaggactctgcagtctattactgtgcaagaggctactat 4266
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTACG 664

Db   4267 tactattctttggactactggggtcaaggaacctcagtcaccgtctc 4313
             |||  |  || ||   ||||||   ||| ||| | ||||||||||
Qy    665 GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT   10
ID   Q43844 standard; DNA; 11529 BP.
AC   Q43844;
DT   20-OCT-1993  (first entry)
DE   Plasmid pAH4602.
KW   Polymerase chain reaction; primer; PCR; amplify; murine;
KW   heavy;light; chain; variable; constant; region; anti-human;
KW   transferrin; receptor; antibody; brain; capillary;
KW   endothelial cell; conjugate; neuropharmaceutical;
KW   diagnostic; agent; tumour; AIDS; stroke; epilepsy;
KW   Parkinsons disease; Alzheimers disease; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   unsure          2070..3769
FT                   /*tag= a
FT   cds             3907..4611
FT                   /*tag= b
FT                   /note= "VH and CH DNA derived by PCR"
FT   signal_peptide  3097..3963
FT                   /*tag= c
FT                   /note= "Leader sequence"
FT   mat_peptide     3964..4611
FT                   /*tag= d
FT                   /note= "Mature VH/CH region"
FT   cds             5000..5044
FT                   /*tag= e
FT                   /note= "Undefined ORF"
FT   cds             5163..5492
FT                   /*tag= f
FT                   /note= "Undefined ORF2"
FT   cds             5589..5909
FT                   /*tag= g
FT                   /note= "Undefined ORF2"
FT   unsure          6118..7321
FT                   /*tag= h
FT   unsure          8177..8201
FT                   /*tag= i
PN   WO9310819-A.
PD   10-JUN-1993.
PF   24-NOV-1992; U10206.
PR   26-NOV-1991; US-800458.
PA   (ALKE-) ALKERMES INC.
PI   Friden PM;
DR   WPI; 93-196742/24.
DR   P-PSDB; R41682-85.
PT   Antibody conjugates specific for transferrin receptor - used
PT   for diagnosis and treatment of cancer, AIDS and neurological
PT   disorders
PS   Disclosure; Fig 11A-G; 151pp; English.
CC   This sequence represents the expression vector pAH4602.
CC   This vector contains open reading frames encoding the heavy
CC   chain variable region (VH) of the antibody 128.1, an
CC   ampicillin resistance gene and a histidine (histidinol)
CC   selection marker.  Transcription of the VH gene is from the
CC   VH promoter of the murine 27.44 gene.  The vector also
CC   includes a heavy chain immunoglobulin enhancer and the human
CC   gammal constant region (CH).  The VH region of 128.1 was
CC   isolated by polymerase chain reaction and cloned into
CC   plasmid pAH4274.  This was achieved by digesting the plasmid
CC   and the product with EcoRV and NheI.  The VH gene was
CC   inserted in-frame with the human gammal CH region CH at the
CC   3' end of the VH-J region by means of a NheI site.  128.1 is
CC   an anti-human transferrin receptor antibody which binds to
CC   the transferrin receptor on brain capillary endothelial
CC   cells.  This antibody may be used in a conjugate in which it
CC   is linked to a neuropharmaceutical or diagnostic agent.  The
CC   conjugate may be used to treat or prevent neurological
CC   disorders eg. brain tumours, AIDS, stroke, epilepsy,
CC   Parkinsons and Alzheimers disease.  It may also be used for
CC   diagnostic methods.
SQ   Sequence   11529 BP;    2175 A;    2217 C;    2148 G;    2058 T;

  Query Match           73.9%;  Score 257;  DB 7;  Length 11529;
  Best Local Similarity  87.0%;  Pred. No. 6.39e-160;
  Matches   302;  Conservative    0;  Mismatches  45;  Indels   0;  Gaps   0;

Db   3967 gtccagctgcaacagtctggacctgaactggtgaagcctggagcttcaatgaagatttcc 4026
            || ||||||||  |||| ||||||||  ||||||||||||||| ||||||||||| |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db   4027 tgcaaggcttctggttactcattcactggctacaccatgaactgggtgaagcagagccat 4086
            |||||| |||||||||||||||||||||| |||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db   4087 ggagagaaccttgagtggattggacgtattaatcctcacaatggtggtactgactacaac 4146
            ||| |||||||||||||||||||| |||||||||| ||||||||| |||| ||||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db   4147 cagaagttcaaggacaaggccccttttaactgtagacaagtcatccaacacagcctacatg 4206
            |||||||||||||| ||||||| | || |||||||||||||| |||| ||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db   4207 gagctcctcagtctgacatctgaggactctgcagtctattactgtgcaagaggctactat 4266
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTACG 664

Db   4267 tactattctttggactactggggtcaaggaacctcagtcaccgtctc 4313
             |||  |  || ||   ||||||   ||| ||| | ||||||||||
Qy    665 GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT   11
ID   Q43847 standard; DNA; 12132 BP.
AC   Q43847;
DT   20-OCT-1993  (first entry)
DE   Plasmid pAH4807.
KW   Polymerase chain reaction; primer; PCR; amplify; murine; pAH4625;
KW   heavy; light; chain; variable; constant; region; anti-human; pAH4807;
KW   transferrin; receptor; antibody; brain; capillary; pAH4808; gamma-2;
KW   endothelial cell; conjugate; neuropharmaceutical; gamma-3; gamma-4;
KW   diagnostic; agent; tumour; AIDS; stroke; epilepsy; monoclonal;
KW   Parkinsons disease; Alzheimers disease; SP2/0 cell; ss.
OS   Synthetic.
```

```
FH   Key             Location/Qualifiers
FT   unsure          11..1710
FT                   /*tag= a
FT   cds             1848..2546
FT                   /*tag= b
FT                   /note= "128.1 VH/human gamma-3 CH1"
FT   cds             2939..2980
FT                   /*tag= c
FT                   /note= "Undefined ORF1"
FT   cds             3133..3177
FT                   /*tag= d
FT                   /note= "Undefined ORF2"
FT   cds             3672..4001
FT                   /*tag= e
FT                   /note= "Undefined ORF3"
FT   cds             4099..4419
FT                   /*tag= f
FT                   /note= "Undefined ORF4"
FT   unsure          4630..5559
FT                   /*tag= g
FT   unsure          5566..5864
FT                   /*tag= h
FT   unsure          6720..6744
FT                   /*tag= i
PN   WO9310819-A.
PD   10-JUN-1993.
PF   24-NOV-1992; U10206.
PR   26-NOV-1991; US-800458.
PA   (ALKE-) ALKERMES INC.
PI   Friden PM;
DR   WPI; 93-196742/24.
DR   P-PSDB; R41710-14.
PT   Antibody conjugates specific for transferrin receptor - used
PT   for diagnosis and treatment of cancer, AIDS and neurological
PT   disorders
PS   Disclosure; Fig 18A-F; 151pp; English.
CC   The sequences given in Q43846-48 represent the expression vectors
CC   pAH4625, pAH4807 and pAH4808.  These vectors represent the cloning
CC   of the different human gamma isotypes, gamma-2, gamma-3 and gamma-4
CC   respectively, with the variable region of the murine monoclonal
CC   antibody 128.1.  These plasmids each encode a chimeric monoclonal
CC   antibody in which the heavy chain (VH) is derived from a murine
CC   source and the sequences encoding CH1, CH2 and CH3 are derived from
CC   a human source.  These vectors in combination with the chimeric light
CC   chain vector, pAG4611 (see also Q43845), were transfected into SP2/0
CC   cells and clones were isolated.  Antibody analysis using biosynthe-
CC   tically labelled proteins, immunoprecipitation and SDS-PAGE indicated
CC   appropriate bands for the heavy and light chains as well as the
CC   assembled antibody for the gamma-3 and gamma-4 chimeras.  No detectable
CC   protein was isolated for the gamma-2 transfectants.  128.1 is an anti-
CC   human transferrin receptor antibody which binds to the transferrin
CC   receptor on brain capillary endothelial cells.  This antibody may be
CC   used in a conjugate in which it is linked to a neuropharmaceutical or
CC   diagnostic agent.  The conjugate may be used to treat or prevent
CC   neurological disorders eg. brain tumours, AIDS, stroke, epilepsy,
CC   Parkinsons and Alzheimers disease.  It may also be used for diagnostic
CC   methods.
SQ   Sequence   12132 BP;    2292 A;    2441 C;    2285 G;    2156 T;

  Query Match           73.9%;  Score 257;  DB 7;  Length 12132;
  Best Local Similarity  87.0%;  Pred. No. 6.39e-160;
  Matches   302;  Conservative    0;  Mismatches  45;  Indels   0;  Gaps   0;

Db   1908 gtccagctgcaacagtctggacctgaactggtgaagcctggagcttcaatgaagatttcc 1967
          || ||||||||  |||| ||||||||  ||||||||||||||| |||||||||||| |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db   1968 tgcaaggcttctggttactcattcactggctacaccatgaactgggtgaagcagagccat 2027
          |||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db   2028 ggagagaaccttgagtggattggacgtattaatcctcacaatggtggtactgactacaac 2087
          ||| ||||||||||||||||||||| |||||||||| ||||||||| |||| ||||||||

Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db   2088 cagaagttcaaggacaaggccccttttaactgtagacaagtcatccaacacagcctacatg 2147
          |||||||||||||| ||||||| | || ||||||||||||||| |||| ||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db   2148 gagctcctcagtctgacatctgaggactctgcagtctattactgtgcaagaggctactat 2207
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTACG 664

Db   2208 tactattctttggactactggggtcaaggaacctcagtcaccgtctc 2254
          |||  |  || ||   ||||||   ||| ||| | ||||||||||
Qy    665 GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT   12
ID   Q45428 standard; cDNA; 443 BP.
AC   Q45428;
DT   17-NOV-1994  (first entry)
DE   KM-750 heavy chain.
KW   Monoclonal antibody; Ab; ganglioside GM2; chimera;
KW   chimeric antibody; expression vector; heavy; light; chain;
KW   hypervariable region; CDR; constant region; hybridoma;
KW   Ig; immunoglobulin; KM-796; KM-750; KM-603; cancer; ss.
OS   Mus musculus.
FH   Key             Location/Qualifiers
FT   cds             33..443
FT                   /*tag= a
FT                   /note= "excluding stop codon"
FT   signal_peptide  33..89
FT                   /*tag= b
FT                   /note= "sig_peptide"
FT   misc_feature    180..194
FT                   /*tag= c
FT                   /note= "hypervariable region 1"
FT   misc_feature    252..287
FT                   /*tag= d
FT                   /note= "hypervariable region 2"
FT   misc_feature    384..410
FT                   /*tag= e
FT                   /note= "hypervariable region 3"
PN   AU9346181-A.
PD   17-MAR-1994.
PF   07-SEP-1993; 046181.
PR   07-SEP-1992; JP-238452.
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
PI   Hanai N,  Hasegawa M,  Koike M,  Kuwana Y,  Nakamura K;
PI   Shitara K;
DR   WPI; 94-126857/16.
DR   P-PSDB; R53330.
PT   Humanised antibody specific for ganglioside GM2 - used for
PT   producing a cytocidal effect on cancers such as melanoma,
PT   neuroblastoma and glioma.
PS   Disclosure; Page 107-108; 191pp; English.
CC   Chimeric human Ab expression vectors are constructed by inserting
CC   the Ab heavy and light chain variable region-encoding cDNA
CC   isolated from hybridomas producing a mouse or rat monoclonal Ab
CC   reacting with the ganglioside GM2 respectively into an expression
CC   vector for use in animal cells which contains the human Ab heavy and
CC   light chain constant region-encoding cDNA.  The expression vectors
CC   are introduced into animal cells and the transformant thus obtained
CC   is cultured for the prodn. of a chimeric human Ab reacting with the
CC   ganglioside GM2.  In contrast to mouse monoclonal Abs, the chimeric
CC   human Abs will not cause anti-mouse Ig Ab prodn. in the patient's
CC   body but show a prolonged blood half-life, with a reduced frequency
CC   of adverse effects, so that it can be expected to be superior
CC   to mouse monoclonal Abs in the efficacy in the treatment of human
CC   cancer, for instance.
CC   Mouse anti-GM2 monoclonal Ab KM-796 and KM-750 and rat KM-603
CC   heavy and light chain sequences are given in Q45426-30.
CC   CDR regions for use in chimeric Abs are indicated in the
CC   Features Table.
SQ   Sequence   443 BP;    109 A;    112 C;    119 G;    103 T;
```

```
  Query Match           71.6%;  Score 249;  DB 11;  Length 443;
  Best Local Similarity  85.9%;  Pred. No. 3.21e-154;
  Matches   298;  Conservative    0;  Mismatches  49;  Indels   0;  Gaps   0;

Db     93 gtccagctgcagcagtctggacctgagctggtgaagcctggggcttcagtgaagatatcc 152
          || ||||||||| |||| ||||||||| |||||||||||||| | |||| ||||||||||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db    153 tgcaaggcttctggatacacattcactgactacaacatggactgggtgaagcagagccat 212
          |||||| ||||||| ||| ||||||||| | ||| |||| |||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    213 ggaaagagccttgagtggattggatatatttatcctaacaatggtggtactggctacaac 272
          ||||||| |||||||||||||||  |||| ||||| |||||||||  ||||  |||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    273 cagaagttcaagagcaaggccacattgactgtagacaagtcctccagcacagcctacatg 332
          |||||||||||| |||||||||||||| |||||||||||||| ||||||||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    333 gagctccacagcctgacatctgaggactctgcagtctattactgtgcaagagcggggagg 392
          ||||||| ||| ||||||||||||||||||||||||||||||||||||    ||  | |
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTACG 664

Db    393 tattactacgcctgggactggggccaagggactctggtcactgtctc 439
          | |  |||  |   | |||||||  ||||||   ||||||  |||||
Qy    665 GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT   13
ID   Q28522 standard; cDNA to mRNA; 356 BP.
AC   Q28522;
DT   23-FEB-1993  (first entry)
DE   Hypercalcaemia agent cDNA portion 1.
KW   Antihuman parathyroid hormone-related protein; monoclonal antibody;
KW   variable region; rodent/human chimeric MAb; constant region; PTHrP; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   cds             7..354
FT                   /*tag=  a
PN   J04228089-A.
PD   18-AUG-1992.
PF   15-MAY-1991; 110565.
PR   15-MAY-1990; JP-124581.
PA   (KANF ) KANEKA CORP.
DR   WPI; 92-320987/39.
DR   P-PSDB; R27010.
PT   Treatment and preventive agent for hypercalcaemia - contg. one of
PT   anti-human para-thyroid-hormone-related protein monoclonal antibody,
PT   a rodent or chimera monoclonal antibody, fused gene and cell
PT   line, etc.
PS   Disclosure; Page 14; 18pp; Japanese.
CC   The sequences given in Q28522-23 encode fragments which were used in
CC   the construction of an agent for the treating and prevention of
CC   hypercalcaemia.  The agent contained a portion of the antihuman
CC   parathyroid hormone-related protein monoclonal antibody (antihuman
CC   PTHrP MAb).  The MAb was used as the active component in the agent.
CC   The agent further comprises a rodent/human chimeric MAb which has a
CC   rodent variable region and a human constant region and recognises
CC   human PTHrP.
SQ   Sequence   356 BP;    87 A;    77 C;    103 G;    89 T;

  Query Match           70.4%;  Score 245;  DB 5;  Length 356;
  Best Local Similarity  85.9%;  Pred. No. 2.27e-151;
  Matches   293;  Conservative    0;  Mismatches  48;  Indels   0;  Gaps   0;

Db     13 gttcagctgcagcagtctggacctgagctggtgaagcctggggcttcagtgaagatatcc 72
          || ||||||||| |||| ||||||||| |||||||||||||| | |||| ||||||||||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     73 tgcaaggcttctggttactcatttactggctactttatgaactgggtgatgcagagccat 132
          |||||| |||||||||||||||| |||||| ||   |||||||||||||| ||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    133 ggaaagagccttgagtggattggacgtattaatccttacaatggtgatactttctacaac 192
          ||||||| |||||||||||||||| ||||||||||||||||||||||||  |||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    193 cagaagttcaagggcaaggccacattgactgtagacaaatcctctagcacagcccacatg 252
          |||||||||||||||||||||||||| |||||||||| || || ||||||||| |||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    253 gagctccggagcctggcatctgaggactctgcagtctattattgtgcaagagggggggtt 312
          |||||||  || ||| |||||||||||||||||||||||||| ||||||||  |||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTACG 664

Db    313 acgacggagtttggttactggggtccagggaccacggtcac 353
              | | || | |  ||||||  ||||||||||||||||
Qy    665 GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCAC 705


RESULT   14
ID   T86309 standard; DNA; 717 BP.
AC   T86309;
DT   06-APR-1998  (first entry)
DE   Single chain anti-disialoganglioside GD2 antibody 5F11-scFv.
KW   Antibody construct; disialoganglioside; GD2; single chain Fv fragment;
KW   scFv; tumour; neuroblastoma; osteosarcoma; soft tissue sarcoma;
KW   tissue imaging; target delivery; toxin; streptavidin;
KW   pro-drug converting enzyme; GD2-targeted lymphocyte; ss.
OS   Synthetic.
PN   WO9734634-A1.
PD   25-SEP-1997.
PF   20-MAR-1997; U04427.
PR   20-MAR-1996; US-013703.
PA   (SLOK ) SLOAN KETTERING INST CANCER RES.
PI   Cheung NV, Guo H, Larson SM, Rivlin K, Sadelain M;
DR   WPI; 97-479996/44.
PT   Recombinant single chain anti-disialoganglioside GD2 antibody -
PT   useful to detect tumour cells expressing GD2 and to target
PT   therapeutic agents, e.g. toxins, to such cells
PS   Disclosure; Page 11; 31pp; English.
CC   The present sequence encodes a recombinant single chain peptide,
CC   5F11-scFv. The peptide is an antibody construct comprising the variable
CC   regions of the heavy and light chains of an antibody against
CC   disialoganglioside (GD2) as a single chain Fv fragment (scFv). GD2 occurs
CC   in many tumours types including neuroblastoma, osteosarcomas and other
CC   soft tissue sarcomas, medulloblastomas, high grade astrocytomas,
CC   melanomas and small cell lung cancer. The peptide can be detectably
CC   labelled, preferably with 99m-Tc, for tissue imaging of cells expressing
CC   GD2. It can also be used to target delivery of a therapeutic or
CC   pre-therapeutic agent, such as a toxin, streptavidin or a pro-drug
CC   converting enzyme, to cells expressing GD2. The peptide may further
CC   comprise CD8 to facilitate the formation of GD2-targeted lymphocytes.
CC   T cells containing the peptide can also be used to target GD2-producing
CC   tumour cells.
SQ   Sequence   717 BP;    180 A;    181 C;    195 G;    159 T;

  Query Match           69.5%;  Score 242;  DB 37;  Length 717;
  Best Local Similarity  86.7%;  Pred. No. 3.10e-149;
  Matches   301;  Conservative    0;  Mismatches  43;  Indels   3;  Gaps   2;

Db      4 gtgaaactgcagcagtcaggacctgaactggtgnagcctggggcttcagtgaagatatcc 63
          ||| | |||||| |||||||||||||  ||||| ||||||| | |||| ||||||||||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     64 tgcaagacttctggtanacaaattcactgaatacaccatgcactgggtgaagcagagccat 123
          |||||||||||||  ||  ||||||||  |||||||| ||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    124 ggaaagagccttgagtggattggaggtattaatcctaacaatggtggtactaactacaag 183
          ||||||| |||||||||||||||  |||||||||| |||||||||  ||||||||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    184 cagaagttcaagggcaaggccacattgactgtagacaagtcctccagcacagcctacatg 243
```

```
          |||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    244 gagctccgcagcctgacatctgaggattctgcagtctattactgtgcaagagatactacg 303
          ||||||| ||| |||||||||||||| |||||||||||||||||||||||      ||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTACG 664

Db    304 gtcccgt--tt-gcttactgggtccaagggaccacggtcaccgtctc 347
          | |  ||  || | |  ||||| |  ||||||||||||||||||||
Qy    665 GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


RESULT   15
ID   T86311 standard; DNA; 1175 BP.
AC   T86311;
DT   06-APR-1998  (first entry)
DE   Single chain anti-disialoganglioside GD2 antibody 5F11-scFv-streptavidin.
KW   Antibody construct; disialoganglioside; GD2; single chain Fv fragment;
KW   scFv; tumour; neuroblastoma; osteosarcoma; soft tissue sarcoma;
KW   tissue imaging; target delivery; toxin; streptavidin;
KW   pro-drug converting enzyme; GD2-targeted lymphocyte; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    1..716
FT                   /*tag=  a
FT                   /note= "encodes single chain Fv fragment of antibody
FT                    5F11"
FT   misc_feature    717..740
FT                   /*tag=  b
FT                   /note= "linker sequence"
FT   misc_feature    741..1175
FT                   /*tag=  c
FT                   /note= "encodes streptavidin"
PN   WO9734634-A1.
PD   25-SEP-1997.
PF   20-MAR-1997; U04427.
PR   20-MAR-1996; US-013703.
PA   (SLOK ) SLOAN KETTERING INST CANCER RES.
PI   Cheung NV, Guo H, Larson SM, Rivlin K, Sadelain M;
DR   WPI; 97-479996/44.
PT   Recombinant single chain anti-disialoganglioside GD2 antibody -
PT   useful to detect tumour cells expressing GD2 and to target
PT   therapeutic agents, e.g. toxins, to such cells
PS   Disclosure; Page 12; 31pp; English.
CC   The present sequence encodes a recombinant single chain peptide,
CC   5F11-scFv-streptavidin. The peptide is an antibody construct comprising
CC   the variable regions of the heavy and light chains of an antibody against
CC   disialoganglioside (GD2) as a single chain Fv fragment (scFv). GD2 occurs
CC   in many tumours types including neuroblastoma, osteosarcomas and other
CC   soft tissue sarcomas, medulloblastomas, high grade astrocytomas,
CC   melanomas and small cell lung cancer. The peptide can be detectably
CC   labelled, preferably with 99m-Tc, for tissue imaging of cells expressing
CC   GD2. It can also be used to target delivery of a therapeutic or
CC   pre-therapeutic agent, such as a toxin, streptavidin (e.g. present
CC   sequence) or a pro-drug converting enzyme, to cells expressing GD2. The
CC   peptide may further comprise CD8 to facilitate the formation of
CC   GD2-targeted lymphocytes. T cells containing the peptide can also be used
CC   to target GD2-producing tumour cells.
SQ   Sequence   1175 BP;    276 A;    346 C;    330 G;    221 T;

  Query Match           69.5%;  Score 242;  DB 37;  Length 1175;
  Best Local Similarity  86.7%;  Pred. No. 3.10e-149;
  Matches   301;  Conservative    0;  Mismatches  43;  Indels   3;  Gaps   2;

Db      4 gtgaaactgcagcagtcaggacctgaactggtgnagcctggggcttcagtgaaagatatcc 63
          ||| | |||||| |||||||||||||  ||||| ||||||| | |||| ||||||||||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     64 tgcaagacttctgganacaaattcactgaatacaccatgcactgggtgaagcagagccat 123
          |||||||||||||  ||  ||||||||   |||||||| ||||||||||||||||||||
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    124 ggaaagagccttgagtggattggaggtattaatcctaacaatggtggtactaactacaag 183
          ||||||| |||||||||||||||  |||||||||| |||||||| ||||||||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    184 cagaagttcaagggcaaggccacattgactgtagacaagtcctccagcacagcctacatg 243
          ||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    244 gagctccgcagcctgacatctgaggattctgcagtctattactgtgcaagagatactacg 303
          ||||||| ||| |||||||||||||| |||||||||||||||||||||||      ||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAGGAGGGTTACG 664

Db    304 gtcccgt--tt-gcttactgggtccaagggaccacggtcaccgtctc 347
          | |  ||  || | |  ||||| |  ||||||||||||||||||||
Qy    665 GACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTC 711


Search completed: Sat Jan  9 16:27:40 1999
Job time : 70 secs.
```

****

(TM)

****

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Jan 9 15:24:53 1999; MasPar time 491.43 Seconds
1268.112 Million cell updates/sec

Tabular output not generated.

Title: >US-08-704-178-1
Description: (364-711) from US08704178.seq (3 of 4)
Perfect Score: 348
N.A. Sequence: 364 TGTGCAGCTGCAGGAGTCAG.....GGGACCACGGTCACCGTCTC 711
Comp: ACACGTCGACGTCCTCAGTC.....CCCTGGTGCCAGTGGCAGAG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl-est55
1:em_est1 2:em_gss1 3:em_gss2 4:em_gss3
Database: genbank-est107
5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est2 17:gb_est20
18:gb_est21 19:gb_est3 20:gb_est4 21:gb_est5 22:gb_est6
23:gb_est7 24:gb_est8 25:gb_est9 26:gb_gss1 27:gb_gss2
28:gb_gss3 29:gb_gss4

Statistics: Mean 9.946; Variance 1.816; scale 5.477

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
	1	160	46.0	310	18	AI180569	uc70f09.r1 Soares mous	9.99e-288
	2	146	42.0	294	15	AI007196	ua73g04.r1 Soares 2NbM	1.51e-256
	3	110	31.6	387	9	AA569186	nm30d10.s1 NCI_CGAP_Li	1.76e-177
	4	85	24.4	279	24	AA300571	EST13661 Testis tumor	5.47e-124
	5	82	23.6	379	8	AA291381	zt44g02.r1 Soares ovar	1.10e-117
c	6	81	23.3	209	16	T28938	EST61186 Homo sapiens	1.37e-115
	7	79	22.7	116	17	AI116236	uc18e08.r1 Soares mous	2.06e-111
	8	76	21.8	431	11	AA710970	vt93h09.r1 Soares mous	3.58e-105
	9	71	20.4	287	16	T27609	EST101034 Homo sapiens	7.44e-95
	10	68	19.5	336	15	AA987559	or83g01.s1 NCI_CGAP_Lu	1.02e-88
	11	66	19.0	265	25	AA377128	EST89660 Small intesti	1.19e-84
	12	66	19.0	291	25	AA377074	EST89603 Small intesti	1.19e-84
c	13	64	18.4	488	15	AA948436	on67a01.s1 Soares_NFL_	1.32e-80
	14	62	17.8	238	16	T29670	EST89669 Homo sapiens	1.39e-76
c	15	59	17.0	252	12	AA754459	97SN1787 Rice Immature	1.35e-70
c	16	57	16.4	247	14	AA897116	am08b07.s1 Soares_NFL_	1.23e-66
	17	56	16.1	630	14	AA895088	vy32h01.r1 Stratagene	1.15e-64
	18	56	16.1	823	23	AA170256	ms87g10.r1 Soares mous	1.15e-64
	19	55	15.8	372	11	AA581192	nd38b11.r1 NCI_CGAP_Co	1.05e-62
	20	54	15.5	266	10	AA402547	zu47h07.r1 Soares ovar	9.45e-61
	21	53	15.2	385	16	R86288	yp12e09.r1 Homo sapien	8.35e-59
	22	49	14.1	252	12	AA754459	97SN1787 Rice Immature	4.19e-51
	23	47	13.5	498	28	AQ180882	HS_3225_A2_D05_MR CIT	2.61e-47
	24	43	12.4	198	6	AA383972	EST97425 Thymus II Hom	7.52e-40
	25	41	11.8	330	21	SSC1D10	S.scrofa mRNA; express	3.42e-36
	26	39	11.2	230	10	AA428970	zv49b02.r1 Soares ovar	1.37e-32
	27	39	11.2	238	8	AA505044	aa63g01.r1 NCI_CGAP_GC	1.37e-32
	28	39	11.2	299	16	T27727	EST13874 Homo sapiens	1.37e-32
c	29	39	11.2	378	12	AA740786	ob32e11.s1 NCI_CGAP_Ki	1.37e-32
	30	38	10.9	136	18	AI181594	uc63d07.r1 Soares mous	8.23e-31
	31	38	10.9	258	25	AA360195	EST69374 Lymph node I	8.23e-31
	32	36	10.3	247	12	AA754458	97SN1784 Rice Immature	2.66e-27
c	33	36	10.3	247	12	AA754458	97SN1784 Rice Immature	2.66e-27
	34	36	10.3	415	9	AA575218	vh19a01.r1 Soares mous	2.66e-27
	35	36	10.3	597	7	AA472093	vh10a05.r1 Soares mous	2.66e-27
	36	35	10.1	303	16	R83139	yp11g03.r1 Homo sapien	1.43e-25
	37	35	10.1	419	19	H73816	ys11b01.r1 Homo sapien	1.43e-25
c	38	35	10.1	474	18	AI147237	qb36f02.x1 Soares_preg	1.43e-25
c	39	35	10.1	485	15	AA988730	or79h12.s1 NCI_CGAP_Lu	1.43e-25
c	40	35	10.1	619	26	FR0004509	F.rubripes GSS sequenc	1.43e-25
	41	34	9.8	408	16	H42300	yo63g08.r1 Homo sapien	7.33e-24
	42	34	9.8	478	16	H43753	yp21g01.r1 Homo sapien	7.33e-24
	43	34	9.8	480	10	AA632998	nq09a07.s1 NCI_CGAP_Lu	7.33e-24
	44	33	9.5	273	24	AA300982	EST13889 Testis tumor	3.59e-22
	45	33	9.5	297	16	T27868	EST18962 Homo sapiens	3.59e-22

ALIGNMENTS

RESULT 1
LOCUS AI180569 310 bp mRNA EST 08-OCT-1998
DEFINITION uc70f09.r1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone 1431017 5' similar to gb:M18512 IG HEAVY CHAIN PRECURSOR V-I REGION (HUMAN); gb:M12376 Mouse immunoglobulin H-chain V-region pseudogene mRNA, complete (MOUSE);, mRNA sequence.
ACCESSION AI180569
NID g3731207
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 310)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:915085
Seq primer: -28m13 rev2 ET from Amersham.
FEATURES Location/Qualifiers
source 1. .310
/organism="Mus musculus"
/strain="C57BL/6J"

/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."
/db_xref="taxon:10090"
/clone="1431017"
/clone_lib="Soares mouse mammary gland NbMMG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"

BASE COUNT 80 a 77 c 85 g 67 t 1 others
ORIGIN

Query Match 46.0%; Score 160; DB 18; Length 310;
Best Local Similarity 79.3%; Pred. No. 9.99e-288;
Matches 230; Conservative 0; Mismatches 58; Indels 2; Gaps 2;

```
Db      4 GTCCAACTGCAGCAGCCTGGGG-TAAGCTTGTGAAGCCTGG-GGTTCAGTGAAGCTGTCC 61
          || || |||||| || | ||   | || | ||||||||||| |||||| ||||| | |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     62 TGCAAGGCTTCTGGCTACACCTTCACCAGCTACTGGATGCAGTGGGTAAAACAGAGGCCT 121
          |||||| ||||||| ||| | |||||  || ||   ||| | ||||| || |||||| | |
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    122 GGACAGGGCATTGAGTGGATCGGAGAGATTGATCCTTCTGATAGCTATACTAACTACAAT 181
          ||| ||  | ||||||||| |||   ||| ||||||   || |  ||||||||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db    182 CAAAAGTTCAAGGGCAAGGCAGCATTGACTGTAGACACATCCTCCAGCACAGCCTACATG 241
          || ||||||||||||||||  |||| ||||||||||  || ||||||||||||||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    242 CAGCTCAGCAGCCTGACATCTGAGGACTCTGCGGTCTATGACTGTGCAAG 291
          |||||  ||| |||||||||||||||||||| |||||| ||||||||||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654
```

RESULT 2
LOCUS AI007196 294 bp mRNA EST 12-JUN-1998
DEFINITION ua73g04.r1 Soares 2NbMT Mus musculus cDNA clone 1363158 5' similar to gb:M18512 IG HEAVY CHAIN PRECURSOR V-I REGION (HUMAN); gb:M12376 Mouse immunoglobulin H-chain V-region pseudogene mRNA, complete (MOUSE);, mRNA sequence.
ACCESSION AI007196
NID g3216753
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 294)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:896378
Trace considered overall poor quality
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .294
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
/db_xref="taxon:10090"
/clone="1363158"
/clone_lib="Soares 2NbMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"

BASE COUNT 85 a 68 c 77 g 64 t
ORIGIN

Query Match 42.0%; Score 146; DB 15; Length 294;
Best Local Similarity 76.7%; Pred. No. 1.51e-256;
Matches 224; Conservative 0; Mismatches 66; Indels 2; Gaps 2;

```
Db      2 GTCCAACTGCAGCAGCCTGGGGCTGAGCATGGTGAGGACTGGGTCTTCAGTGAAGCTGTA 61
          || || |||||| || | ||  |||||  |||||| | ||||   |||| ||||| | |
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGG-TGGTGAAGCCTGGAGGTTCAATGAAGATATC 423

Db     62 CTGCAAGGATTCAGGCTACACCTTCACCAGCTACTGGATGCATTGGGTGAAGCAGAGGCA 121
          |||||||  ||| || ||| | |||||  || ||   ||| | ||||||||||||||| ||
Qy    424 CTGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCA 483

Db    122 TATACAAGGCATTGAATGGATTGGTAACATTGACCCTTCAGATAGTGGAACTCACTACAA 181
          |  | |   | |||| ||||||||    ||| | ||||   || |||  ||| |||||||
Qy    484 TGGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAA 543

Db    182 TCAAAAGTTCAAGGACAAGGCCACATTGACTGTAGACAAATTCACCAGCACAGCCTACAG 241
          || |||||||||| |||||||||||| |||||||||| |   ||||||||||||||||
Qy    544 CCAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACA- 602

Db    242 TGCAGCTCAGCAGACTGACATCTGAGGACTCTGCGGTCTATTACTGTGCAAG 293
          || |||||  ||| ||||||||||||||||||| ||||||||||||||||||
Qy    603 TGGAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654
```

RESULT 3
LOCUS AA569186 387 bp mRNA EST 09-SEP-1997
DEFINITION nm30d10.s1 NCI_CGAP_Lip2 Homo sapiens cDNA clone IMAGE:1061683 similar to gb:L02325 IG HEAVY CHAIN PRECURSOR V-I REGION (HUMAN);, mRNA sequence.
ACCESSION AA569186
NID g2342240
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 387)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

```
  JOURNAL   Unpublished (1997)
COMMENT
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Insert Length: 1722   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 284.
FEATURES             Location/Qualifiers
     source          1. .387
                     /organism="Homo sapiens"
                     /note="Vector: pAMP10; mRNA made from liposarcoma, cDNA
                     made by oligo-dT priming. Non- directionally cloned.
                     Size-selected on agarose gel, average insert  size 600 bp.
                     Reference:  Krizman et al. (1996) Cancer Research
                     56:5380-5383."
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1061683"
                     /clone_lib="NCI_CGAP_Lip2"
                     /tissue_type="liposarcoma"
                     /lab_host="DH10B"
     mRNA            <1. .>387
BASE COUNT       89 a     95 c    118 g     85 t
ORIGIN

  Query Match           31.6%;  Score 110;  DB 9;  Length 387;
  Best Local Similarity  69.0%;  Pred. No. 1.76e-177;
  Matches   200;  Conservative    0;  Mismatches  90;  Indels   0;  Gaps   0;

Db      98 GTCCAACTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGGGGCCTCAGTGAAGCTTTCC 157
           || || |||  | |||| ||  ||||||||  |||||||||  |  ||| ||||| | |||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     158 TGTAAGGCATCTGGATTCACCCTCACCAACTGCCATATGCACTGGGTGCGACAGGCCCCT 217
           || ||| | ||||| | | |  ||||   |  |   ||| ||||||||   |||  || |
Qy     425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db     218 GGGCAAGGGCTTGAGTGGGTGGGAATGATCAATTCTAGTGATGGTTATATAAGTAACGCA 277
           ||  |    |||||||||  | ||| | ||  ||  |||||  |||  |   ||
Qy     485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db     278 CAGAAGTTCCAGGGCAGAGTCACCATGACCAGGGACACGTCCACGAGCACAGTCTATATG 337
           |||||||||  ||||||  | |||  | ||    |||| |||  | |||||||  ||| |||
Qy     545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db     338 GAGCTTAGCAGCCTGAGATCTGAGGACACGGCCGTCTATTACTGTGCGAG 387
           |||||   ||| |||| |||||||||| | || ||||||||||||||| ||
Qy     605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654


RESULT    4
LOCUS       AA300571      279 bp    mRNA            EST       18-APR-1997
DEFINITION  EST13661 Testis tumor Homo sapiens cDNA 5' end similar to similar
            to immunoglobulin mu heavy chain, VDJC regions (GB:M18512), mRNA
            sequence.
ACCESSION   AA300571
NID         g1953132
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 279)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: THC166571
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .279
                     /organism="Homo sapiens"
                     /note="Organ: testis; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
                     /db_xref="ATCC (inhost):192212"
                     /db_xref="taxon:9606"
                     /clone_lib="Testis tumor"
                     /sex="male"
                     /dev_stage="adult"
     mRNA            <1. .>279
BASE COUNT       63 a     74 c     85 g     56 t      1 others
ORIGIN

  Query Match           24.4%;  Score 85;  DB 24;  Length 279;
  Best Local Similarity  72.3%;  Pred. No. 5.47e-124;
  Matches   136;  Conservative    0;  Mismatches  52;  Indels   0;  Gaps   0;

Db      92 GTGCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGGGGCCTCAGTGAAGGTCTCC 151
           |||||||||  | |||| ||  ||||||||  |||||||||  |  ||| ||||| | |||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     152 TGCAAGGCTTCTGGATACACCTTCACCGGCTACTATATGCACTGGGTGCGACAGGCCCCT 211
           |||||| ||||||| ||| | |||| ||| ||   ||| ||||||||   |||  || |
Qy     425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db     212 GGACAAGGGCTTNAGTGGATGGGATGGATCAACCCTAAAAGTGGTGGCACAAACTATGCA 271
           ||| |   ||| ||||||| |||   || || ||| | | |||||  || |||||
Qy     485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db     272 CAGAAGTT 279
           ||||||||
Qy     545 CAGAAGTT 552


RESULT    5
LOCUS       AA291381      379 bp    mRNA            EST       08-AUG-1997
DEFINITION  zt44g02.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 725234
            5' similar to gb:M18512 IG HEAVY CHAIN PRECURSOR V-I REGION
            (HUMAN);, mRNA sequence.
```

```
ACCESSION   AA291381
NID         g1939359
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 379)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1576   Std Error: 0.00
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 361.
FEATURES              Location/Qualifiers
     source           1. .379
                      /organism="Homo sapiens"
                      /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library constructed by Bento Soares and
                      M.Fatima Bonaldo."
                      /db_xref="taxon:9606"
                      /clone="725234"
                      /clone_lib="Soares ovary tumor NbHOT"
                      /sex="Female"
                      /tissue_type="ovarian tumor"
                      /lab_host="DH10B (ampicillin resistant)"
     mRNA             <1. .>379
BASE COUNT       85 a     82 c    121 g     91 t
ORIGIN

  Query Match           23.6%;  Score 82;  DB 8;  Length 379;
  Best Local Similarity  65.9%;  Pred. No. 1.10e-117;
  Matches   191;  Conservative    0;  Mismatches  97;  Indels   2;  Gaps   2;

Db      37 GTTCAGTTGGTGCAGTCTGGACCTGAGGTGAAGAAGC-TGG-GGCCCAGTGAGGGTCTCC 94
           || ||| ||  | |||| ||||||||||||  ||||| ||| ||  || ||| | | |||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db      95 TGTAAGACCTCTGGTTACACCTTTGGCACTTTTGGTATCAACTGGGTGCGGCAGGCCCCT 154
           || ||||| |||||||| | ||             || |||||||||  ||||  || |
Qy     425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db     155 GGACAAGGGCTTGAGTGGATGGGATGGATCAGCGTTTACAATGGTAAGACAAACTTTGCA 214
           ||| |    |||||||||| |||   || |    |||||||||| | || ||||
Qy     485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db     215 GGGAGTTTTCAGGACAGAGTCAGCTTGACCACTGACACATCCACGAATACAGCCTACATG 274
             ||  ||  ||| ||  | ||  || ||    ||||  ||  | |  ||||||||||||
Qy     545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db     275 GAGCTGAGGAACCTGAGATCTGACGACACGGCCATATATTACTGTGCGAG 324
           |||||    |  |||| |||||| ||| | ||  | ||||||||||| ||
Qy     605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654
```

```
RESULT    6
LOCUS       T28938        209 bp    mRNA            EST       06-SEP-1995
DEFINITION  EST61186 Homo sapiens cDNA 5' end similar to immunoglobulin mu
            heavy chain, VDJ regions (GB:M17751) (HT:3055).
ACCESSION   T28938
NID         g611036
KEYWORDS    EST.
SOURCE      human primer-M13 Reverse library-Human White blood cells.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 209)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W.,
            Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            FitzGerald,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghagen,N.S.M.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle Jr,P.S.,
            Kelley,J.M., Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M.,
            Merrick,J.M., Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T.,
            Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L.,
            Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R.,
            Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A.,
            Coleman,T.A., Collins,E.-J., Dimke,D., Feng,P., Ferrie,A.,
            Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M.,
            Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H.,
            Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C.,
            Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A.,
            Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
  TITLE     Initial Assessment of Human Gene Diversity and Expression Patterns
            Based Upon 52 Million Basepairs of cDNA Sequence
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Venter, JC
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: tdbinfo@tdb.tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please contact the TIGR Database
            (tdbinfo@tdb.tigr.org).
FEATURES              Location/Qualifiers
     source           1. .209
                      /organism="Homo sapiens"
     mRNA             <1. .>209
BASE COUNT       43 a     62 c     48 g     55 t      1 others
ORIGIN

  Query Match           23.3%;  Score 81;  DB 16;  Length 209;
  Best Local Similarity  71.7%;  Pred. No. 1.37e-115;
  Matches   134;  Conservative    0;  Mismatches  53;  Indels   0;  Gaps   0;

Db      15 GTAATGGTGACTCTGCCCTGGAACTTCTGTAAATATTTTGTGTTACCATTGCCAGCGTTG 74
           ||||  ||| |  |||||| |||||||||     || || || | |||||||  ||  ||
Cp     573 GTAAATGTGGCCTTGCCCTTGAACTTCTGGTTGTAGTTAGTATCACCATTGTAAGGATTA 514

Db      75 ATCCATCCCATCCACTCAAGCCTTTGTCCGGGGGCCTGGCGCACCCAATGCATAGCATAG 134
           ||   ||| ||||||||||||  | | |||  ||  |||   |||||| | ||| |  | |
Cp     513 ATAAGTCCAATCCACTCAAGGTTCTTTCCATGGCTCTGCTTCACCCAGTTCATGGTGTGG 454

Db     135 CTAGTGAAGGTGTATCCAGAAGCCTTGCAGGAAACCTTCACTGAGGCCCCAGGCTTCTTC 194
           | |||||| | ||| ||||||| |||||||| | ||||| |||  | |||||||||  |
Cp     453 CCAGTGAATGAGTAACCAGAAGTCTTGCAGGATATCTTCATTGAACCTCCAGGCTTCACC 394

Db     195 ACCTCAG 201
           |||||||
Cp     393 ACCTCAG 387
```

```
RESULT    7
LOCUS       AI116236      116 bp    mRNA            EST       02-SEP-1998
DEFINITION  uc18e08.r1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone
            1398374 5' similar to gb:M87789 IG GAMMA-1 CHAIN C REGION (HUMAN);
            gb:X67210 M.musculus rearranged immunoglobulin gamma 2b heavy
            (MOUSE);, mRNA sequence.
ACCESSION   AI116236
NID         g3516560
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 116)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:910090
            Trace considered overall poor quality
            Seq primer: -28ml3 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .116
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Organ: mammary gland; Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                     oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                     constructed and normalized by Bento Soares and M.Fatima
                     Bonaldo."
                     /db_xref="taxon:10090"
                     /clone="1398374"
                     /clone_lib="Soares mouse mammary gland NbMMG"
                     /sex="male"
                     /tissue_type="mammary gland"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
BASE COUNT       38 a     32 c     25 g     21 t
ORIGIN

  Query Match           22.7%;  Score 79;  DB 17;  Length 116;
  Best Local Similarity  86.2%;  Pred. No. 2.06e-111;
  Matches    94;  Conservative    0;  Mismatches  15;  Indels   0;  Gaps   0;

Db       2 AGAAGTTCAAGAACAAGGCCACACTGACTGCAGACAAATCCTCCAGCACAGCCTACATGC 61
           |||||||||||  |||||||||| | |||| |||||| || ||||||||||||||||||
Qy     546 AGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATGG 605

Db      62 AACTCAGCAGCCTGACATCTGAGGACACTGCGGTCTATTACTGTGGAAG 110
           | |||  ||| ||||||||||||||| |||| |||||||||||| |||
Qy     606 AGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654
```

```
RESULT    8
LOCUS       AA710970      431 bp    mRNA            EST       24-DEC-1997
DEFINITION  vt93h09.r1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone
            1178753 5' similar to gb:M18512 IG HEAVY CHAIN PRECURSOR V-I REGION
            (HUMAN); gb:X70423 M.musculus mRNA for monoclonal antibody heavy
            chain gamma (MOUSE);, mRNA sequence.
ACCESSION   AA710970
NID         g2720888
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 431)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:636601
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 198.
FEATURES             Location/Qualifiers
     source          1. .431
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Organ: mammary gland; Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                     oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                     constructed and normalized by Bento Soares and M.Fatima
                     Bonaldo."
                     /db_xref="taxon:10090"
                     /clone="1178753"
                     /clone_lib="Soares mouse mammary gland NbMMG"
                     /sex="male"
                     /tissue_type="mammary gland"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
BASE COUNT      112 a    103 c    112 g    104 t
ORIGIN

  Query Match           21.8%;  Score 76;  DB 11;  Length 431;
  Best Local Similarity  63.9%;  Pred. No. 3.58e-105;
  Matches   175;  Conservative    0;  Mismatches  99;  Indels   0;  Gaps   0;

Db     124 AGTCTGGACCTGAGCTGAAGAAGCCTGGAGAGACAGTCAAGATCTCCTGCAAGGCTTCTG 183
           |||| |||||||||| ||  |||||||||||    || | ||||| |||||||||| ||||||
Qy     378 AGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCCTGCAAGACTTCTG 437

Db     184 GGTATACCTTCACAACCTATGGAATGAGCTGGGTGAAACAGGCTCCAGGAAAGGGTTTAA 243
           | ||  | |||||   | |    |||| ||||||||| |||   |  ||||||    |
Qy     438 GTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCATGGAAAGAACCTTG 497

Db     244 AGTGGATGGGCTGGATAAACACCTACTCTGGAGTGCCAACATATGCTGATGACTTCAAGG 303
           ||||||| ||    || ||  | |||  ||| |   | |  ||     |  | |||||||
```

```
Qy    498 AGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAACCAGAAGTTCAAGG 557

Db    304 GACGGTTTGCCTTCTCTTTGGAAACCTCTGCCAGCACTGCCTATTTGCAGATCAACAACC 363
          |   |    | ||  || | || |  ||  |||||||| |||||  || || ||  ||  |
Qy    558 GCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATGGAGCTCCTCAGTC 617

Db    364 TCAAAAATGAGGACACGGCTACATATTTCTGTGC 397
          | | |  |||||||  | ||    |||| ||||||
Qy    618 TGACATCTGAGGACTCTGCAGTCTATTACTGTGC 651


RESULT    9
LOCUS       T27609         287 bp    mRNA            EST       06-SEP-1995
DEFINITION  EST101034 Homo sapiens cDNA 5' end similar to immunoglobulin mu
            heavy chain, VDJC regions (GB:M18512) (HT:3056).
ACCESSION   T27609
NID         g609707
KEYWORDS    EST.
SOURCE      human primer-M13 Reverse library-Human Pancreas.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 287)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W.,
            Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            FitzGerald,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghagen,N.S.M.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle Jr,P.S.,
            Kelley,J.M., Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M.,
            Merrick,J.M., Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T.,
            Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L.,
            Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R.,
            Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A.,
            Coleman,T.A., Collins,E.-J., Dimke,D., Feng,P., Ferrie,A.,
            Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M.,
            Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H.,
            Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C.,
            Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A.,
            Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
  TITLE     Initial Assessment of Human Gene Diversity and Expression Patterns
            Based Upon 52 Million Basepairs of cDNA Sequence
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Venter, JC
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: tdbinfo@tdb.tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please contact the TIGR Database
            (tdbinfo@tdb.tigr.org).
FEATURES              Location/Qualifiers
     source           1. .287
                      /organism="Homo sapiens"
     mRNA             <1. .>287
BASE COUNT       62 a     77 c     85 g     60 t      3 others
ORIGIN

  Query Match           20.4%;  Score 71;  DB 16;  Length 287;
  Best Local Similarity  70.3%;  Pred. No. 7.44e-95;
  Matches   123;  Conservative    0;  Mismatches  52;  Indels   0;  Gaps   0;

Db     74 GTGCACCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGGGGCCTCAGTGAAGGTCTCC 133
          ||||| |||  | |||| ||  ||||||||  |||||||| | |||  |||| | |||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db    134 TGCAAGACTTCTGGATACGCCTTCACCGACTACTATATACACTGGATTCGACAGGCCCCT 193
          |||||||||||||| ||| | ||||| | | ||   ||  |||||  |    |||  || |
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    194 GGACAAGGGCTTGAATGGATGGGATGGATGAACCCTCGCAGTGGTGCCACAAACT 248
          ||| |    ||||| ||||| |||    || || |||  || |||||  || ||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACT 539


RESULT   10
LOCUS       AA987559       336 bp    mRNA            EST       23-JUL-1998
DEFINITION  or83g01.s1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1602480 3'
            similar to gb:M87789 IG GAMMA-1 CHAIN C REGION (HUMAN);, mRNA
            sequence.
ACCESSION   AA987559
NID         g3172923
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 336)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Insert Length: 801   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 322.
FEATURES              Location/Qualifiers
     source           1. .336
                      /organism="Homo sapiens"
                      /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                      modified polylinker; 1st strand cDNA was prepared from
                      neuroendocrine lung carcinoid, and was then primed with a
                      Not I - oligo(dT) primer. Double-stranded cDNA was ligated
                      to Eco RI adaptors (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of the modified
                      pT7T3 vector.  Library is normalized. Library was
                      constructed by Bento Soares and M. Fatima Bonaldo."
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1602480"
                      /clone_lib="NCI_CGAP_Lu5"
                      /tissue_type="carcinoid"
                      /lab_host="DH10B"
BASE COUNT       82 a     86 c    100 g     68 t
ORIGIN

  Query Match           19.5%;  Score 68;  DB 15;  Length 336;
  Best Local Similarity  61.7%;  Pred. No. 1.02e-88;
  Matches   179;  Conservative    0;  Mismatches 111;  Indels   0;  Gaps   0;

Db     32 GTCCAGATAGAACAGTCAGGGGCTGAGGTGAAGAAGCCAGGGTCTTCGGTGAAGGTCACC 91
          || ||| |  |  |||||||  ||||||||  |||||| ||   |||  ||||| |  ||
Qy    365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     92 TGCTTGGCCTCTGGAGGCACCTTCGGTAGTTCGACTGTCAGCTGGGTGCGACAGGCCCCT 151
          |||  | | |||||   | | |||  | |    ||  | | |||||||   |||  || |
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    152 GGACAACGGCCTGAGTGGATGGGGAAGATCATCCCCATCGCTCATTCAGCGAACTACGCA 211
          ||| |    | |||||||||  ||    || |  ||   |  | |   | ||||||
Qy    485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544
```

```
Db    212 CCGAAGTTTCAGGACAGAATCACAATAACCGCGGACAAATCTACGGCGTCAGTCTACATG 271
          | ||||||  ||| ||    |||| | || |  ||||| ||  |     ||| |||||||
Qy    545 CAGAAGTTCAAGGGCAAGGCCACATTTACTGTAGACAAGTCGTCCAGCACAGCCTACATG 604

Db    272 GAATTGACCAGCCTGATATTTGAGGACACGGCCATATATTATTGTGCGAG 321
          ||  |   ||| |||| || ||||||| | ||  | ||||| ||||| ||
Qy    605 GAGCTCCTCAGTCTGACATCTGAGGACTCTGCAGTCTATTACTGTGCAAG 654


RESULT   11
LOCUS       AA377128      265 bp    mRNA            EST       21-APR-1997
DEFINITION  EST89660 Small intestine I Homo sapiens cDNA 5' end similar to
            similar to immunoglobulin heavy chain, VDJ region, mRNA sequence.
ACCESSION   AA377128
NID         g2029456
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 265)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: THC167579
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES              Location/Qualifiers
     source           1. .265
                      /organism="Homo sapiens"
                      /note="Organ: small intestine; Vector: pBluescript SK-;
                      Site_1: EcoRI; Site_2: XhoI"
                      /db_xref="ATCC (inhost):181546"
                      /db_xref="taxon:9606"
                      /clone_lib="Small intestine I"
                      /dev_stage="adult"
     mRNA             <1. .>265
BASE COUNT       57 a     69 c     79 g     58 t      2 others
ORIGIN

  Query Match           19.0%;  Score 66;  DB 25;  Length 265;
  Best Local Similarity  69.8%;  Pred. No. 1.19e-84;
  Matches   113;  Conservative    0;  Mismatches  49;  Indels   0;  Gaps   0;

Db    101 CAGCTGGTGCAGTCTGGAGGTGAGGTGAAGAAGCCTGGGGCCTCANTGAAGGTCTCCTGC 160
          ||||||  | |||| |||  |||||||  |||||||||  |  |||  |||||  | ||||||
Qy    368 CAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCCTGC 427

Db    161 AAGANTTCTGGTTACACCTTTACCACATACGGTATCAGCTGGGTGCGACAGGCCCCTGGA 220
          |||| |||||||||| | || ||      ||   || | |||||||   |||  || ||||
Qy    428 AAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCATGGA 487

Db    221 CAAGGGCTCGAGTGGATGGGATGGATCACCGGTTACAATGGT 262
            |    || |||||||| |||   || |    ||||||||||
Qy    488 AAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGT 529


RESULT   12
LOCUS       AA377074      291 bp    mRNA            EST       21-APR-1997
DEFINITION  EST89603 Small intestine I Homo sapiens cDNA 5' end similar to
            similar to immunoglobulin heavy chain, VDJ regions (GB:Z14165),
            mRNA sequence.
ACCESSION   AA377074
NID         g2029413
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 291)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES              Location/Qualifiers
     source           1. .291
                      /organism="Homo sapiens"
                      /note="Organ: small intestine; Vector: pBluescript SK-;
                      Site_1: EcoRI; Site_2: XhoI"
                      /db_xref="ATCC (inhost):181500"
                      /db_xref="taxon:9606"
                      /clone_lib="Small intestine I"
                      /dev_stage="adult"
     mRNA             <1. .>291
BASE COUNT       67 a     73 c     89 g     61 t      1 others
ORIGIN
```

```
  Query Match           19.0%;  Score 66;  DB 25;  Length 291;
  Best Local Similarity  66.0%;  Pred. No. 1.19e-84;
  Matches   134;  Conservative    0;  Mismatches  69;  Indels   0;  Gaps   0;

Db      75 GTGCAGCTGGTGCAATCTGGGGCTGAGGTGAAGGAGCCTGGGTCTTCTGTGAAAGTNTCG 134
           |||||||||  | | || ||  ||||||||  | |||||||   |||  ||||  | ||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     135 TGCAAGGCTATTGGATACACCTTCACTGATTATACTATCAACTGGGTGCGACAGGCCCCT 194
           ||||||  ||  |||  |||  | |||||||   | ||  ||  |||||||||   |||  || |
Qy     425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db     195 GGACAGGGACTTGAGTGGATGGGATGGATGAACCCAGCAAATGGAGGCTCCGACTGTGCA 254
           |||  ||   ||||||||||| |||   ||  ||  ||    |||||  |   |  |||
Qy     485 GGAAAGAACCTTGAGTGGATTGGACTTATTAATCCTTACAATGGTGATACTAACTACAAC 544

Db     255 CAGAAGTTTCGGGGGAGAGTCAC 277
           ||||||||   |||  |  | |||
Qy     545 CAGAAGTTCAAGGGCAAGGCCAC 567


RESULT   13
LOCUS       AA948436      488 bp    mRNA            EST       23-JUN-1998
DEFINITION  on67a01.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:1561704 3' similar to gb:M18512 IG HEAVY CHAIN PRECURSOR V-I
            REGION (HUMAN);, mRNA sequence.
ACCESSION   AA948436
NID         g3109689
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 488)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1240    Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 148.
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Homo sapiens"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1561704"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
BASE COUNT      108 a    111 c     88 g    181 t
ORIGIN

  Query Match           18.4%;  Score 64;  DB 15;  Length 488;
  Best Local Similarity  71.4%;  Pred. No. 1.32e-80;
  Matches   115;  Conservative    0;  Mismatches  45;  Indels   1;  Gaps   1;

Db     271 CTCGCACAGTAATACACGGCCATGTCCTCAGATCTCAGGCTGCTCAGCTCCGTGTAGGCT 330
           || |||||||||| || ||   |||||||||| |||| |||   |||||| ||||||||
Cp     654 CTTGCACAGTAATAGACTGCAGAGTCCTCAGATGTCAGACTGAGGAGCTCCATGTAGGCT 595

Db     331 GTGCTCATGGACGTGTCCCTGGTTATGGTGACTCTGCCCTGGAACTTCTGTGCATATGTT 390
           |||||    ||| ||||    || |  ||| |  |||||| |||||||||    ||  |
Cp     594 GTGCTGGACGACTTGTCTACAGTAAATGTGGCCTTGCCCTTGAACTTCTGGTTGTAGTTA 535

Db     391 GTGTTACCATTGCCAGCATTGATCCACCCA-TCCACTCAAG 430
           || | |||||||  || ||| ||    ||| ||||||||||
Cp     534 GTATCACCATTGTAAGGATTAATAAGTCCAATCCACTCAAG 494


RESULT   14
LOCUS       T29670        238 bp    mRNA            EST       06-SEP-1995
DEFINITION  EST89669 Homo sapiens cDNA 5' end similar to immunoglobulin heavy
            chain V region (GB:X61012) (HT:3230).
ACCESSION   T29670
NID         g611768
KEYWORDS    EST.
SOURCE      human primer=M13 Reverse library=Human Small intestine.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 238)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W.,
            Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            FitzGerald,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghagen,N.S.M.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle Jr,P.S.,
            Kelley,J.M., Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M.,
            Merrick,J.M., Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T.,
            Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L.,
            Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R.,
            Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A.,
            Coleman,T.A., Collins,E.-J., Dimke,D., Feng,P., Ferrie,A.,
            Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M.,
            Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H.,
            Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C.,
            Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A.,
            Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
  TITLE     Initial Assessment of Human Gene Diversity and Expression Patterns
            Based Upon 52 Million Basepairs of cDNA Sequence
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Venter, JC
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: tdbinfo@tdb.tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please contact the TIGR Database
            (tdbinfo@tdb.tigr.org).
FEATURES             Location/Qualifiers
     source          1. .238
                     /organism="Homo sapiens"
     mRNA            <1. .>238
BASE COUNT       46 a     70 c     69 g     52 t      1 others
ORIGIN

  Query Match           17.8%;  Score 62;  DB 16;  Length 238;
  Best Local Similarity  72.8%;  Pred. No. 1.39e-76;
  Matches    99;  Conservative    0;  Mismatches  37;  Indels   0;  Gaps   0;

Db     103 GTTCAGCTGGTGCAGTCTGGAGCTGAGGTGAAGAAGCCTGGGGCCTCAGTGAAGGTCTCC 162
           || ||||||  | |||| ||| ||||||||  |||||||||  | |||  ||||| | |||
Qy     365 GTGCAGCTGCAGGAGTCAGGACCTGAGGTGGTGAAGCCTGGAGGTTCAATGAAGATATCC 424

Db     163 TGCAAGGCTTCTGGTTACACCTTTACCAGCTACGGTATCAGCTGGGTGCGACAGGCCCCT 222
```

```
          |||||| |||||||||||| | || ||  || ||   || | |||||||   |||  || |
Qy    425 TGCAAGACTTCTGGTTACTCATTCACTGGCCACACCATGAACTGGGTGAAGCAGAGCCAT 484

Db    223 GGACAAGGGCTTGAGT 238
          ||| |    |||||||
Qy    485 GGAAAGAACCTTGAGT 500


RESULT   15
LOCUS       AA754459      252 bp    mRNA            EST       20-JAN-1998
DEFINITION  97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
            cDNA clone 97SN1787, mRNA sequence.
ACCESSION   AA754459
NID         g2801165
KEYWORDS    EST.
SOURCE      rice.
  ORGANISM  Oryza sativa
            Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
            Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
            Poales; Poaceae; Oryza.
REFERENCE   1  (bases 1 to 252)
  AUTHORS   Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R, Moon,E.P.,
            Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
            Lee,M.C. and Eun,M.Y.
  TITLE     Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
  JOURNAL   Unpublished (1998)
COMMENT
            Contact: Eun M.Y.
            Department of Cytogenetics
            National Inst. of Agri. Sci. and Tech, RDA
            Suwon, Kyunggido, Korea
            Tel: 82 331 290 0301
            Fax: 82 331 290 0307
            Email: myeun@sun20.asti.re.kr
            Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
            University, Yongin, Korea. 449-728 bhnahm@bioserver.myongji.ac.kr
            Seq primer: M13 Reverse Primer.
FEATURES             Location/Qualifiers
     source          1. .252
                     /organism="Oryza sativa"
                     /cultivar="Milyang23"
                     /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
                     XhoI; Directional cDNA library inserted into lambda ZAPII
                     vector  at 5'end with EcoRI and 3' end with Xho I site."
                     /db_xref="taxon:4530"
                     /clone="97SN1787"
                     /clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
                     /tissue_type="Immature Seed"
                     /dev_stage="5 days after pollination"
                     /lab_host="E. coli SOLR"
BASE COUNT        5 a      21 c      12 g      35 t     179 others
ORIGIN

  Query Match           17.0%;  Score 59;  DB 12;  Length 252;
  Best Local Similarity  11.1%;  Pred. No. 1.35e-70;
  Matches    26;  Conservative  120;  Mismatches  86;  Indels   3;  Gaps   2;

Db     21 TSYBCHGNBVWVCVASHGNYMSVHNCTBRGTHCDCKNVNWSTMTWGTVNWBNVSGDWHYW 80
          |:::|:   :::   ::  :::::   ::  :| |: : :: : :||: :  :: : : :
Cp    691 TGCGCCCCAGACATCGAAGTACCAGTCCGTAACCCTCCTTGCACAGTAATAGACTGCAGA 632

Db     81 BVBNTKVDVGNHTRCSRWRBVTRMAHYHDYTNCBBYNNNDYHMWHBBMYBBTGCMTCTMW 140
          : : | :::  : :   :: : :  :::::   :::   :::  ::::::: |  |  :
Cp    631 GTCCTCAGATGTCAGACTGAGGAGCTCCATGTAGGCTGTGCTGGACGACTTGTCTACAGT 572

Db    141 CWBHYNTKCTASGWHTSTNYDVKSSTNTWGVTBSYDKSMHGYWCSBBVKYHTKVSTTRAT 200
           : :   :|   | :  |  ::: :|  : :| ::: : : ::|: : ::: :: ||:||
Cp    571 AAATGTGGCCTTGCCCTTGAACTTCTGGTTGTAGTTAGTATCACCATTGTAAGGATTAAT 512

Db    201 R-SYTCVRKYCVMWMTKKV--VKKYHVVBBGCHBTDSKCKTMWMTNKHVMTSTTD 252
          : :: |::::|::::  ::  ::::::::::||::|:::   : :  ::::|: |:
Cp    511 AAGTCCAATCCACTCAAGGTTCTTTCCATGGCTCTGCTTCACCCAGTTCATGGTG 457
```

Search completed: Sat Jan 9 15:33:13 1999
Job time : 500 secs.

```
**************************************************************************
 ___    ___   _____   ____    _____   ____    _____   _    _
|  \  /  /| |  ___| |  __|  |  ___| |  __|  |  ___| | |  | |
|   \/   | | |___  | |__   | |___  | |__   | |___  | |__| |
| |\  /| | |___  | |  __|  |___  | |  __|  |___  | |  __  |
| | \/ | | |___| | | |___   ___| | | |___   ___| | | |  | |
|_|    |_| |_____| |____|  |_____| |____|  |_____| |_|  |_|  (TM)
**************************************************************************

        Release 3.1A John F. Collins, Biocomputing Research Unit.
        Copyright (c) 1993-1998  University of Edinburgh, U.K.
              Distribution rights by Oxford Molecular Ltd

MPsrch_nn   n.a. -  n.a. database search, using Smith-Waterman algorithm

Run on:          Sat Jan  9 17:10:52 1999;  MasPar time 100.47 Seconds
                                            921.250 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-704-178-1
Description:     (322-363) from US08704178.seq (4 of 4)
Perfect Score:   42
N.A. Sequence:      322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363
     Comp:              TCCAAGATGGAGACCAAGACCATTTAGAAGACTTCCATTTCC

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        567134 seqs, 1101898692 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        embl55
                   1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
                   7:em_om 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
                   13:em_ro 14:em_vi
Database:        genbank107
                   15:gb_ba1 16:gb_ba2 17:gb_htg 18:gb_in 19:gb_om 20:gb_ov
                   21:gb_pat 22:gb_ph 23:gb_pl1 24:gb_pl2 25:gb_pr1
                   26:gb_pr2 27:gb_pr3 28:gb_ro 29:gb_st 30:gb_sts 31:gb_sy
                   32:gb_un 33:gb_vi

Statistics:      Mean 7.713;  Variance 4.202;  scale 1.836

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	42	100.0	711	21	I32406	Sequence 1 from patent	3.22e-13
2	42	100.0	711	31	S39590	anti-erbB2 imnmunotoxi	3.22e-13
3	38	90.5	720	21	I32407	Sequence 2 from patent	1.36e-10
4	34	81.0	846	28	MUSALC	Mus musculus (clone pC	5.00e-08
5	34	81.0	855	28	MUSALCA	Mus musculus (clone pC	5.00e-08
6	31	73.8	47	21	I32414	Sequence 9 from patent	3.74e-06
c 7	30	71.4	47	21	I32413	Sequence 8 from patent	1.54e-05
8	24	57.1	717	31	SYNBET2X	Synthetic single-chain	5.35e-02
9	24	57.1	743	31	SYN7A41	Synthetic single-chain	5.35e-02
10	24	57.1	744	31	SYNDA44X	Synthetic single-chain	5.35e-02
c 11	20	47.6	228	23	S82816	ndhA=NDH-A [Hordeum vu	8.14e+00
c 12	20	47.6	590	23	HVNDHA590	Hordeum vulgare chloro	8.14e+00
c 13	20	47.6	1636	24	HVU010977	Hordeum vulgare chloro	8.14e+00
c 14	20	47.6	7368	33	HPU31780	Human papillomavirus t	8.14e+00
15	20	47.6	13316	16	AE001306	Chlamydia trachomatis	8.14e+00
16	20	47.6	100808	26	HS94G16	Human DNA sequence fro	8.14e+00
17	19	45.2	250	30	G31000	human STS SHGC-18931,	2.67e+01
18	19	45.2	607	19	BTIGA48B	B.taurus cosmid-derive	2.67e+01
c 19	19	45.2	1532	28	RNU40627	Rattus norvegicus clon	2.67e+01
c 20	19	45.2	1864	23	SCYKL180W	S.cerevisiae chromosom	2.67e+01
c 21	19	45.2	1876	28	RNU40628	Rattus norvegicus clon	2.67e+01
c 22	19	45.2	9202	33	VLVGAGA	Visna virus proviral g	2.67e+01
c 23	19	45.2	9202	33	VLVCGA	Visna lentivirus, Icel	2.67e+01
c 24	19	45.2	9202	21	A15114	Visna virus RNA genome	2.67e+01
c 25	19	45.2	9202	33	S55323	gag. . .rev [visna vir	2.67e+01
c 26	19	45.2	9203	33	VLVCG	Visna lentivirus compl	2.67e+01
c 27	19	45.2	9221	33	VLVLV1A	Visna virus Icelandic	2.67e+01
c 28	19	45.2	9225	33	VLVLV1B	Visna virus Icelandic	2.67e+01
c 29	19	45.2	13232	18	DDID16466	Cellular slime mold Mi	2.67e+01
c 30	19	45.2	15926	24	SPBC25D12	S.pombe chromosome II	2.67e+01
31	19	45.2	36495	24	AB015473	Arabidopsis thaliana g	2.67e+01
c 32	19	45.2	38477	23	SCPEKGA	S.cerevisiae sequence	2.67e+01
33	19	45.2	40007	24	SPCC320	S.pombe chromosome III	2.67e+01
34	19	45.2	56500	26	AB013139	Homo sapiens gene for	2.67e+01
35	19	45.2	65589	27	AF049895	Homo sapiens nibrin (N	2.67e+01
36	19	45.2	81357	27	AC005792	Homo sapiens chromosom	2.67e+01
c 37	19	45.2	83478	23	AB011482	Arabidopsis thaliana g	2.67e+01
c 38	19	45.2	88295	26	AC004612	Homo sapiens PAC clone	2.67e+01
c 39	19	45.2	95643	24	F6N15	Arabidopsis thaliana B	2.67e+01
c 40	19	45.2	97451	23	F19K23	Sequence of BAC F19K23	2.67e+01
41	19	45.2	150613	23	AB001684	Chlorella vulgaris C-2	2.67e+01
c 42	19	45.2	152503	26	HUAC003009	Homo sapiens Chromosom	2.67e+01
c 43	19	45.2	184919	27	AF069291	Homo sapiens BAC clone	2.67e+01
c 44	19	45.2	195142	26	AC004554	Homo sapiens Xp22 BAC	2.67e+01
c 45	19	45.2	216021	26	HUAC004787	Homo sapiens Chromosom	2.67e+01

```
                                ALIGNMENTS

RESULT    1
LOCUS       I32406        711 bp    DNA             PAT       07-JAN-1997
DEFINITION  Sequence 1 from patent US 5587458.
ACCESSION   I32406
NID         g1823197
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 711)
  AUTHORS   King,C.Richter, Kasprzyk,P.G. and Bird,R.E.
  TITLE     Anti-erbB-2 antibodies, combinations thereof, and therapeutic and
            diagnostic uses thereof
  JOURNAL   Patent: US 5587458-A 1 24-DEC-1996;
FEATURES             Location/Qualifiers
     source          1. .711
                     /organism="unknown"
BASE COUNT      175 a    182 c    190 g    164 t
ORIGIN

  Query Match          100.0%;  Score 42;  DB 21;  Length 711;
  Best Local Similarity 100.0%;  Pred. No. 3.22e-13;
  Matches   42;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    321 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 362
          ||||||||||||||||||||||||||||||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT    2
LOCUS       S39590        711 bp                    SYN       10-FEB-1993
DEFINITION  anti-erbB2 imnmunotoxin antigen binding region [mice, Other
            Synthetic Recombinant Partial, 711 nt].
ACCESSION   S39590
NID         g251113
KEYWORDS    .
SOURCE      Mus sp.
```

```
  ORGANISM  Mus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 711)
  AUTHORS   Batra,J.K., Kasprzyk,P.G., Bird,R.E., Pastan,I. and King,C.R.
  TITLE     Recombinant anti-erbB2 immunotoxins containing Pseudomonas exotoxin
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 89 (13), 5867-5871 (1992)
  MEDLINE   92335198
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 108547] from the original journal article.
            This sequence comes from Fig. 1.
FEATURES             Location/Qualifiers
     source          1. .711
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
     gene            1. .711
                     /partial
                     /gene="anti-erbB2 imnmunotoxin antigen binding region"
     CDS             1. .711
                     /partial
                     /gene="anti-erbB2 imnmunotoxin antigen binding region"
                     /note="This sequence comes from Fig. 1."
                     /codon_start=1
                     /db_xref="PID:g251114"
                     /translation="MDLQLTQSPAILSASPGEKVTMTCRATPSVSYMHWYQQKPGSSP
                     KPWIYTTSNLASGVPARFSGGGSGTSYSLTVSRVEAEDAATYYCQQWSRSPPTFGGGS
                     KLEIKGSTSGSGKSSEGKGVQLQESGPEVVKPGGSMKISCKTSGYSFTGHTMNWVKQS
                     HGKNLEWIGLINPYNGDTNYNQKFKGKATFTVDKSSSTAYMELLSLTSEDSAVYYCAR
                     RVTDWYFDVWGAGTTVTVS"
BASE COUNT      175 a    182 c    190 g    164 t
ORIGIN

  Query Match          100.0%;  Score 42;  DB 31;  Length 711;
  Best Local Similarity 100.0%;  Pred. No. 3.22e-13;
  Matches   42;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    321 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 362
          ||||||||||||||||||||||||||||||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT    3
LOCUS       I32407        720 bp    DNA             PAT       07-JAN-1997
DEFINITION  Sequence 2 from patent US 5587458.
ACCESSION   I32407
NID         g1823198
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 720)
  AUTHORS   King,C.Richter, Kasprzyk,P.G. and Bird,R.E.
  TITLE     Anti-erbB-2 antibodies, combinations thereof, and therapeutic and
            diagnostic uses thereof
  JOURNAL   Patent: US 5587458-A 2 24-DEC-1996;
FEATURES             Location/Qualifiers
     source          1. .720
                     /organism="unknown"
BASE COUNT      184 a    178 c    179 g    179 t
ORIGIN

  Query Match          90.5%;  Score 38;  DB 21;  Length 720;
  Best Local Similarity  95.2%;  Pred. No. 1.36e-10;
  Matches   40;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Db    315 AGGTTCTACCTCCGGATCTGGTAAATCTTCTGAAGGTAAAGG 356
          |||||||||||| || ||||||||||||||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT    4
LOCUS       MUSALC        846 bp    mRNA            ROD       11-JUL-1995
DEFINITION  Mus musculus (clone pCT.scFv) antibody mRNA, 3' end of cds.
ACCESSION   L43543
NID         g896287
KEYWORDS    antibody.
SOURCE      Mus musculus (clone: pCT.scFv) cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   Paul,S., Sun,M. and Gao,Q.-S.
  TITLE     Catalysis by natural and hybrid single chain Fv constructs
  JOURNAL   Unpublished (1995)
FEATURES             Location/Qualifiers
     source          1. .846
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="pCT.scFv"
                     /cell_line="c23.5"
                     /cell_type="hybridoma"
     mRNA            <1. .>846
     CDS             <1. .846
                     /note="putative"
                     /codon_start=1
                     /product="antibody"
                     /db_xref="PID:g896290"
                     /translation="DVLMTQTPLTLSVTIGQPASISCKSSQSLLHTDGKTYLIWLLQR
                     PGQSPKRLIYLVSKLDSGVPDRFTGSGSGTDFTLKISRVEAEDLGVYYCWQGTHFPQT
                     FGGGTKLEIKRADAAPGSTSGSGKSSEGKGGVKLVESGGGLVKPGGSLKLSCAASGFT
                     FSIYGMSWFRQTPEKRLEWVATISGGDTYTYYPDSVKGRFTISRDNAKNNLFLQMSSL
                     RSEDTALYFCGRGIAYWGQGTLVTVSATKTTPPSVYPAAAHHHHHHGAAEQKLISEED
                     LNGAA"
     V_region        <1. .354
                     /note="putative"
                     /product="antibody light chain"
     CDS             <1. .>354
                     /note="This CDS feature is included to show the
                     translation of the corresponding V_region. Presently
                     translation qualifiers on V_region features are illegal."
                     /codon_start=1
                     /db_xref="PID:g896288"
                     /translation="DVLMTQTPLTLSVTIGQPASISCKSSQSLLHTDGKTYLIWLLQR
                     PGQSPKRLIYLVSKLDSGVPDRFTGSGSGTDFTLKISRVEAEDLGVYYCWQGTHFPQT
                     FGGGTKLEIKRADAAP"
     misc_feature    355. .396
                     /note="putative"
                     /function="linkers between VL and VH domains"
     V_region        <397. .735
                     /note="putative"
                     /product="antibody heavy chain"
     CDS             <397. .>735
                     /note="This CDS feature is included to show the
                     translation of the corresponding V_region. Presently
                     translation qualifiers on V_region features are illegal."
                     /codon_start=1
                     /db_xref="PID:g896289"
                     /translation="GVKLVESGGGLVKPGGSLKLSCAASGFTFSIYGMSWFRQTPEKR
                     LEWVATISGGDTYTYYPDSVKGRFTISRDNAKNNLFLQMSSLRSEDTALYFCGRGIAY
                     WGQGTLVTVSA"
     misc_feature    736. .765
                     /note="c-myc peptide-1; putative"
                     /function="facilitates recombinant protein identification"
     misc_feature    775. .792
                     /note="poly-histidine; putative"
                     /function="facilitates recombinant protein purification"
BASE COUNT      213 a    203 c    228 g    202 t
ORIGIN

  Query Match          81.0%;  Score 34;  DB 28;  Length 846;
  Best Local Similarity  90.5%;  Pred. No. 5.00e-08;
  Matches   38;  Conservative    0;  Mismatches   4;  Indels   0;  Gaps   0;

Db    354 AGGTTCTACCTCCGGCTCTGGGAAATCGTCTGAAGGTAAAGG 395
```

```
          |||||||||||| || ||||| ||||| ||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT    5
LOCUS       MUSALCA       855 bp    mRNA            ROD       11-JUL-1995
DEFINITION  Mus musculus (clone pCT.kvhd1) antibody mRNA, 3' end of cds.
ACCESSION   L43544
NID         g896291
KEYWORDS    antibody.
SOURCE      Mus musculus cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 855)
  AUTHORS   Paul,S., Sun,M. and Gao,Q.-S.
  TITLE     Catalysis by natural and hybrid single chain Fv constructs
  JOURNAL   Unpublished (1995)
FEATURES             Location/Qualifiers
     source          1. .855
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /cell_type="hybridoma"
     V_region        <1. .354
                     /note="putative"
                     /product="antibody light chain"
     CDS             <1. .855
                     /codon_start=1
                     /product="antibody"
                     /db_xref="PID:g896294"
                     /translation="DVLMTQTPLTLSVTIGQPASISCKSSQSLLHTDGKTYLIWLLQR
                     PGQSPKRLIYLVSKLDSGVPDRFTGSGSGTDFTLKISRVEAEDLGVYYCWQGTHFPQT
                     FGGGTKLEIKRADAAPGSTSGSGKSSEGKGQVQLQESGPGLVAPSQSLSITCTVSGFS
                     LTGYGVNWVRQPPGKGLEWLGMIWGDGNTDYNSALKSRLSISKDNSKSQVFLKMNSLH
                     TDDTARYYCARERDYRLDYWGQGTTVTVSSTKTTPPSVYPAAAHHHHHHGAAEQKLIS
                     EEDLNGAA"
     CDS             <1. .>354
                     /note="This CDS feature is included to show the
                     translation of the corresponding V_region. Presently
                     translation qualifiers on V_region features are illegal."
                     /codon_start=1
                     /db_xref="PID:g896292"
                     /translation="DVLMTQTPLTLSVTIGQPASISCKSSQSLLHTDGKTYLIWLLQR
                     PGQSPKRLIYLVSKLDSGVPDRFTGSGSGTDFTLKISRVEAEDLGVYYCWQGTHFPQT
                     FGGGTKLEIKRADAAP"
     misc_feature    355. .396
                     /note="putative"
                     /function="linkers between VL and VH domains"
     V_region        <397. .744
                     /note="putative"
                     /product="antibody heavy chain"
     CDS             <397. .>744
                     /note="This CDS feature is included to show the
                     translation of the corresponding V_region. Presently
                     translation qualifiers on V_region features are illegal."
                     /codon_start=1
                     /db_xref="PID:g896293"
                     /translation="QVQLQESGPGLVAPSQSLSITCTVSGFSLTGYGVNWVRQPPGKG
                     LEWLGMIWGDGNTDYNSALKSRLSISKDNSKSQVFLKMNSLHTDDTARYYCARERDYR
                     LDYWGQGTTVTVSS"
     misc_feature    745. .774
                     /note="c-myc peptide-1; putative"
                     /function="facilitates recombinant protein identification"
     misc_feature    784. .801
                     /note="poly-histidine; putative"
                     /function="facilitates recombinant protein purification"
BASE COUNT      231 a    217 c    217 g    190 t
ORIGIN

  Query Match           81.0%;  Score 34;  DB 28;  Length 855;
  Best Local Similarity 90.5%;  Pred. No. 5.00e-08;
  Matches    38;  Conservative    0;  Mismatches   4;  Indels   0;  Gaps   0;

Db    354 AGGTTCTACCTCCGGCTCTGGGAAATCGTCTGAAGGTAAAGG 395
          |||||||||||| || ||||| ||||| ||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT    6
LOCUS       I32414        47 bp    DNA             PAT       07-JAN-1997
DEFINITION  Sequence 9 from patent US 5587458.
ACCESSION   I32414
NID         g1823205
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   King,C.Richter, Kasprzyk,P.G. and Bird,R.E.
  TITLE     Anti-erbB-2 antibodies, combinations thereof, and therapeutic and
            diagnostic uses thereof
  JOURNAL   Patent: US 5587458-A 9 24-DEC-1996;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
BASE COUNT       11 a      7 c     16 g     13 t
ORIGIN

  Query Match           73.8%;  Score 31;  DB 21;  Length 47;
  Best Local Similarity 100.0%;  Pred. No. 3.74e-06;
  Matches    31;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 CTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 31
          |||||||||||||||||||||||||||||||
Qy    333 CTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT    7
LOCUS       I32413        47 bp    DNA             PAT       07-JAN-1997
DEFINITION  Sequence 8 from patent US 5587458.
ACCESSION   I32413
NID         g1823204
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   King,C.Richter, Kasprzyk,P.G. and Bird,R.E.
  TITLE     Anti-erbB-2 antibodies, combinations thereof, and therapeutic and
            diagnostic uses thereof
  JOURNAL   Patent: US 5587458-A 8 24-DEC-1996;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
BASE COUNT       15 a      9 c     10 g     13 t
ORIGIN

  Query Match           71.4%;  Score 30;  DB 21;  Length 47;
  Best Local Similarity 100.0%;  Pred. No. 1.54e-05;
  Matches    30;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 GAAGATTTACCAGAACCAGAGGTAGAACCT 30
          ||||||||||||||||||||||||||||||
Cp    351 GAAGATTTACCAGAACCAGAGGTAGAACCT 322


RESULT    8
LOCUS       SYNBET2X      717 bp    mRNA            SYN       01-MAR-1996
DEFINITION  Synthetic single-chain Fv fusion protein (Bet 2/212) mRNA,
            constructed from variable light and heavy chain regions of rat
            monoclonal antibody Bet 2, 3' end of cds.
ACCESSION   L17036
NID         g310725
KEYWORDS    IgM-binding protein; fusion protein; fusion protein Bet 2/212;
            monoclonal antibody.
```

```
SOURCE      Artificial gene cDNA to mRNA.
  ORGANISM  artificial sequence
            artificial sequence.
REFERENCE   1  (bases 1 to 717)
  AUTHORS   Lee,T.K., Rollence,M.L., Hallberg,P.L., Oelkuct,M.S., Dodd,S.W.,
            Nagle,J.W. and Filpula,D.R.
  TITLE     Production of engineered IgM-binding single-chain antibodies in
            Escherichia coli
  JOURNAL   J. Ind. Microbiol. 14 (5), 371-376 (1995)
  MEDLINE   95336669
FEATURES             Location/Qualifiers
     source          1. .717
                     /organism="artificial sequence"
                     /db_xref="taxon:29278"
     CDS             <1. .708
                     /standard_name="Bet 2/212 sFv protein"
                     /codon_start=1
                     /transl_table=11
                     /function="binds mouse IgM"
                     /product="fusion protein"
                     /db_xref="PID:g310726"
                     /translation="DVTQSPSLLSASVGDRVTLSCKGSQNIVNYLAWYQQKLGEAPKL
                     LIFNTNSLQTGIPSRFSGSGSGTDYTLTISGLQPEDVATYFCYQYKNGYTFGAGTKLE
                     LKGSTSGSGKSSEGKGEVQLVESGGGLVQPGSSLKVSCVASRFTFSSYVMHWFRQAPE
                     NGIEWLAYINTDSSSAHYAETVKGRFTISRDNAKNTVDMQLSSLRSEDTAMYFCARGG
                     IKVPVDYWGQGVMVTVS"
     misc_feature    1. .312
                     /note="variable light chain derived from rat monoclonal
                     antibody Bet 2 (ATCC HB88)"
     misc_feature    313. .354
                     /note="linker segment"
     misc_feature    355. .717
                     /note="variable heavy chain derived from rat monoclonal
                     antibody Bet 2 (ATCC HB88)"
BASE COUNT      193 a    170 c    186 g    168 t
ORIGIN

  Query Match           57.1%;  Score 24;  DB 31;  Length 717;
  Best Local Similarity  78.6%;  Pred. No. 5.35e-02;
  Matches    33;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    312 AGGCTCTACTTCCGGTAGCGGCAAATCCTCTGAAGGCAAAGG 353
          ||| ||||| || |||   || ||||| |||||||| |||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT    9
LOCUS       SYN7A41       743 bp    DNA             SYN       04-MAR-1993
DEFINITION  Synthetic single-chain antigen-binding protein gene (7A4-1/212 SCA)
            constructed from the antigen-binding (Fab) fragment of mouse
            catalytic monoclonal antibody NPN43C9 and linker DNA, partial cds.
ACCESSION   M68968
NID         g207839
KEYWORDS    7A41/212 SCA protein; fusion protein; monoclonal antibody;
            single-chain antigen-binding protein.
SOURCE      Mouse DNA and synthetic DNA.
  ORGANISM  artificial sequence
            artificial sequence.
REFERENCE   1  (bases 1 to 743)
  AUTHORS   Gibbs,R.A., Posner,B.A., Filpula,D.R., Dodd,S.W., Finkelman,M.A.
            J., Lee,T.K., Wroble,M., Whitlow,M. and Benkovic,S.J.
  TITLE     Construction and characterization of a single-chain catalytic
            antibody
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 88, 4001-4004 (1991)
  MEDLINE   91219500
FEATURES             Location/Qualifiers
     source          1. .743
                     /organism="artificial sequence"
                     /db_xref="taxon:29278"
     gene            1. .735
                     /gene="7A4-1/212"
     CDS             <1. .735
                     /gene="7A4-1/212"
                     /codon_start=1
                     /transl_table=11
                     /product="7A4-1/212 SCA fusion protein"
                     /db_xref="PID:g207840"
                     /translation="DVVMTQTPSSLAMSVGQKVTMSCKSSQSLLNISNQKNYLAWYQQ
                     KPGQSPKLLVYFASTRESGVPDRFIGSGSGTDFTLTISSVQAEDQADYFCQQHYRAPR
                     TFGGGTKLEIKGSTSGSGKSSEGKGQVQLVESGPGLVAPSQSLSITCTVSGISLSRYN
                     VHWVRQSPGKGLEWLGMIWGGGSIEYNPALKSRLSISKDNSKSQIFLKMNSLQTDDSA
                     MYYCVSYGYGGDRFSYWGQGTLVTVS"
     misc_feature    1. .339
                     /gene="7A4-1/212"
                     /note="light-chain variable region of mouse antibody
                     NPN43C9"
     misc_feature    340. .381
                     /gene="7A4-1/212"
                     /note="synthetic linker"
     misc_feature    382. .735
                     /gene="7A4-1/212"
                     /note="heavy-chain variable region of mouse antibody
                     NPN43C9"
BASE COUNT      195 a    176 c    190 g    182 t
ORIGIN

  Query Match           57.1%;  Score 24;  DB 31;  Length 743;
  Best Local Similarity  78.6%;  Pred. No. 5.35e-02;
  Matches    33;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    339 AGGCTCTACTTCCGGTAGCGGCAAATCCTCTGAAGGCAAAGG 380
          ||| ||||| || |||   || ||||| |||||||| |||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT   10
LOCUS       SYNDA44X      744 bp    mRNA            SYN       01-MAR-1996
DEFINITION  Synthetic single-chain Fv fusion protein (DA4.4/212) mRNA,
            constructed from variable light and heavy chain regions of mouse
            monoclonal antibody DA4.4, 3' end of cds.
ACCESSION   L17037
NID         g310739
KEYWORDS    IgM-binding protein; fusion protein; fusion protein DA4.4/212;
            monoclonal antibody.
SOURCE      Artificial gene cDNA to mRNA.
  ORGANISM  artificial sequence
            artificial sequence.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   Lee,T.K., Rollence,M.L., Hallberg,P.L., Oelkuct,M.S., Dodd,S.W.,
            Nagle,J.W. and Filpula,D.R.
  TITLE     Production of engineered IgM-binding single-chain antibodies in
            Escherichia coli
  JOURNAL   J. Ind. Microbiol. 14 (5), 371-376 (1995)
  MEDLINE   95336669
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="artificial sequence"
                     /db_xref="taxon:29278"
     CDS             <1. .735
                     /standard_name="DA4.4/212 sFv protein"
                     /codon_start=1
                     /transl_table=11
                     /function="binds human IgM"
                     /product="fusion protein"
                     /db_xref="PID:g310740"
                     /translation="DVVMTQSPSSLAMSVGQKVTMSCKSSQSLLNSSNQKNYLAWYQQ
                     KPGQSPELLVYFASTRESGVPDRFIGSGSGTDFTLTISSVQAEDLADYFCQQHYSTPF
                     TFGSGTKLEIKGSTSGSGKSSEGKGQVQLQQPGAEFVKPGAPVKLSCKASGYPFTTYW
                     VNWMKQRPGRGLEWIGRIDPYDSETLYNQKFKDKATLTVDKSSSTAYIQLSSLTSEDS
                     AVYYCARETYDYPFAYWGQGTLVTVS"
     misc_feature    1. .339
                     /note="variable light chain derived from mouse monoclonal
                     antibody DA4.4 (ATCC HB57)"
     misc_feature    340. .381
                     /note="linker segment"
     misc_feature    382. .744
```

```
                /note="variable heavy chain derived from mouse monoclonal
                antibody DA4.4 (ATCC HB57)"
BASE COUNT      192 a    187 c    186 g    179 t
ORIGIN

  Query Match           57.1%;  Score 24;  DB 31;  Length 744;
  Best Local Similarity  78.6%;  Pred. No. 5.35e-02;
  Matches    33;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    339 AGGCTCTACTTCCGGTAGCGGCAAATCCTCTGAAGGCAAAGG 380
          ||| ||||| || |||   || ||||| |||||||| |||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT   11
LOCUS       S82816        228 bp    DNA             PLN       28-DEC-1996
DEFINITION  ndhA=NDH-A [Hordeum vulgare=barley, cv. Hassan, leaf extracts,
            Chloroplast, 228 nt].
ACCESSION   S82816
NID         g1754724
KEYWORDS    .
SOURCE      barley cv. Hassan leaf extracts.
  ORGANISM  Chloroplast Hordeum vulgare
            Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
            Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
            Liliopsida; Poales; Poaceae; Hordeum.
REFERENCE   1  (bases 1 to 228)
  AUTHORS   Martin,M., Casano,L.M. and Sabater,B.
  TITLE     Identification of the product of ndhA gene as a thylakoid protein
            synthesized in response to photooxidative treatment
  JOURNAL   Plant Cell Physiol. 37 (3), 293-298 (1996)
  MEDLINE   96245444
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 179114] from the original journal article.
            This sequence comes from Fig. 1.
FEATURES             Location/Qualifiers
     source          1. .228
                     /organism="Hordeum vulgare"
                     /chloroplast
                     /db_xref="taxon:4513"
     gene            1. .228
                     /partial
                     /note="NDH-A"
                     /gene="ndhA"
     CDS             1. .228
                     /partial
                     /gene="ndhA"
                     /note="This sequence comes from Fig. 1."
                     /codon_start=1
                     /product="NDH-A"
                     /db_xref="PID:g1754725"
                     /translation="PFDLPEAEEELVAGYQTEYSGIKYGLFYLVSYLNLLVSSLFVTV
                     LYLGGWNFSIPYISFFDFFQMNKAVGILEMTM"
BASE COUNT       63 a     35 c     38 g     92 t
ORIGIN

  Query Match           47.6%;  Score 20;  DB 23;  Length 228;
  Best Local Similarity  91.7%;  Pred. No. 8.14e+00;
  Matches    22;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Db      7 GATTTACCAGAAGCAGAGGAAGAA 30
          |||||||||||| |||||| ||||
Cp    348 GATTTACCAGAACCAGAGGTAGAA 325


RESULT   12
LOCUS       HVNDHA590     590 bp    DNA             PLN       09-OCT-1997
DEFINITION  Hordeum vulgare chloroplast DNA for ndhA gene 3'end (exons 1 & 2).
ACCESSION   Y13729
NID         g2239133
KEYWORDS    NDH-A protein; ndhA gene.
SOURCE      Hordeum vulgare.
  ORGANISM  Chloroplast Hordeum vulgare
            Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
            Embryophyta; vascular plants; seed plants; Magnoliophyta;
            Liliopsida; Poales; Poaceae; Hordeum.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   Sabater,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1997) Sabater B., Biologia Vegetal, Fac.
            Ciencias, Campus Universitario, Universidad de Alcala de Henares,
            Cra. Madrid-Barcelona km 33,6, Alcala de Henares. Madrid 28871
            SPAIN
REFERENCE   2  (bases 1 to 590)
  AUTHORS   Lopez,C., Freyer,R., Guera,A., Maier,R.M., Martin,M., Sabater,B.
            and Kossel,H.
  TITLE     Sequence of ndhA gene of barley (Hordeum vulgare L.) plastid.
            Transcript editing in several graminean organs
  JOURNAL   Plant Physiol. 115, 313-313 (1997)
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="Hordeum vulgare"
                     /chloroplast
                     /strain="Hassan"
                     /db_xref="taxon:4513"
                     /tissue_type="leaves"
     gene            13. .551
                     /gene="ndhA"
     CDS             <13. .551
                     /gene="ndhA"
                     /codon_start=3
                     /exception="RNA editing"
                     /evidence=experimental
                     /product="NDH-A protein"
                     /db_xref="PID:e321930"
                     /db_xref="PID:g2239134"
                     /translation="SNSSSTVDIVEAQSKYGFFGWNIWRQPIGFLVFLISSLAECERL
                     PFDLPEAEEELVAGYQTEYSGIKYGLFYLVSYLNLLVSSLFVTVLYLGGWNFSIPYIS
                     FFDFFQMNKAVGILEMTMGIFITLTKAYLFLFISITIRWTLPRMRMDQLLNLGWKFLL
                     PISLGNLLLTTSSQLVSL"
     misc_feature    25
                     /gene="ndhA"
                     /note="C to U RNA editing"
     misc_feature    532
                     /gene="ndhA"
                     /note="C to U RNA editing"
BASE COUNT      180 a     89 c     92 g     229 t
ORIGIN

  Query Match           47.6%;  Score 20;  DB 23;  Length 590;
  Best Local Similarity  91.7%;  Pred. No. 8.14e+00;
  Matches    22;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Db    153 GATTTACCAGAAGCAGAGGAAGAA 176
          |||||||||||| |||||| ||||
Cp    348 GATTTACCAGAACCAGAGGTAGAA 325


RESULT   13
LOCUS       HVU010977    1636 bp    DNA             PLN       08-SEP-1998
DEFINITION  Hordeum vulgare chloroplast ndhA gene, partial exons 1 and 2.
ACCESSION   AJ010977
NID         g3560132
KEYWORDS    NADH dehydrogenase; ndhA gene; subunit A.
SOURCE      Hordeum vulgare.
  ORGANISM  Chloroplast Hordeum vulgare
            Eukaryota; Viridiplantae; Charophyta/Embryophyta group;
            Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta;
            Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
REFERENCE   1  (bases 1 to 1636)
  AUTHORS   Vogel,J., Hess,W.R. and Borner,T.
  TITLE     Precise branch point mapping and quantification of splicing
            intermediates
  JOURNAL   Nucleic Acids Res. 25 (10), 2030-2031 (1997)
  MEDLINE   97277186
REFERENCE   2  (bases 1 to 1636)
```

```
  AUTHORS   Vogel,J., Hess,W.R. and Borner,T.
  TITLE     A comparative analysis of group II introns in chloroplasts of
            higher plants
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1636)
  AUTHORS   Vogel,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-SEP-1998) Vogel J., Department of Biology (Genetics),
            Humboldt University, Chausseestrasse 117, Berlin D-10115, Germany
FEATURES             Location/Qualifiers
     source          1. .1636
                     /organism="Hordeum vulgare"
                     /chloroplast
                     /db_xref="taxon:4513"
     gene            1. .1636
                     /gene="ndhA"
     exon            <1. .333
                     /gene="ndhA"
                     /number=1
     intron          334. .1365
                     /gene="ndhA"
                     /note="subgroup IIB2"
     stem_loop       1290. .1323
                     /gene="ndhA"
                     /note="domain V"
     stem_loop       1328. .1362
                     /gene="ndhA"
                     /note="domain VI"
     misc_feature    1358
                     /gene="ndhA"
                     /note="branch point (bulging A) as determined by
                     lariat-specific RT-PCR"
     exon            1366. .>1636
                     /gene="ndhA"
                     /number=2
BASE COUNT      528 a    252 c    270 g    586 t
ORIGIN

  Query Match           47.6%;  Score 20;  DB 24;  Length 1636;
  Best Local Similarity  91.7%;  Pred. No. 8.14e+00;
  Matches    22;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Db   1506 GATTTACCAGAAGCAGAGGAAGAA 1529
          |||||||||||| |||||| ||||
Cp    348 GATTTACCAGAACCAGAGGTAGAA 325


RESULT   14
LOCUS       HPU31780     7368 bp    DNA             VRL       18-OCT-1995
DEFINITION  Human papillomavirus type 22, complete genome.
ACCESSION   U31780
NID         g1020178
KEYWORDS    .
SOURCE      Human papillomavirus type 22.
  ORGANISM  Human papillomavirus type 22
            Viruses; dsDNA viruses, no RNA stage; Papovaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7368)
  AUTHORS   Delius,H.
  TITLE     Sequenced by Hajo Delius, Deutsches Krebsforschungzentrum,
            Angewandte Tumorvirologie, I.N.F. 506, W-6900 Heidelberg, Germany
  JOURNAL   Unpublished
REFERENCE   2  (sites)
  AUTHORS   Kremsdorf,D., Favre,M., Jablonska,S., Obalek,S., Rueda,L.A.,
            Lutzner,M.A., Blanchet-Bardon,C., Van Voorst Vader,P.C. and Orth,G.
  TITLE     Molecular cloning and characterization of the genomes of nine newly
            recognized human papillomavirus types associated with
            epidermodysplasia verruciformis
  JOURNAL   J. Virol. 52 (3), 1013-1018 (1984)
  MEDLINE   85033930
REFERENCE   3  (bases 1 to 7368)
  AUTHORS   Farmer,A.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-1995) Andrew D. Farmer, HIV Sequence Database,
            Los Alamos National Labroatory, T-10, Mail Stop K710, Los Alamos,
            NM 87501, USA
COMMENT     HPV22 was originally isolated from macules on the chest of an
            Italian epidermodysplasia verruciformis (EV) patient [2]. The HPV22
            genome, like that of HPVs 9, 15, 17a/b, 23, 37, 38, is smaller than
            most PV genomes at approximately 7.4 kb. Phylogenetic
            reconstructions based on DNA sequences of established types
            indicate that HPV22 is most closely related to HPVs 23 and 38, and
            then to 15, 17, 37 and 9. Although Kremsdorf et al [2] found
            substantial cross-hybridization between HPV22 and HPV19, nucleotide
            sequence comparison fails to support a close relationship between
            these two types.
FEATURES             Location/Qualifiers
     source          1. .7368
                     /organism="Human papillomavirus type 22"
                     /note="cloned HPV-22 was obtained from the Papillomavirus
                     Reference Center, Heidelberg and sequenced by Dr. H.
                     Delius"
                     /db_xref="taxon:37954"
     gene            149. .646
                     /gene="E6"
     CDS             149. .646
                     /gene="E6"
                     /note="putative"
                     /codon_start=1
                     /product="transforming protein E6"
                     /db_xref="PID:g1020179"
                     /translation="MQPLVVIYALLAYYLSRMGCYSVFMALQRPLTVQQLSDKLTVPV
                     VDLLLPCRFCSRFLTYLELRQFDYKNLQLIWTDEDFVFACCSGCAYASAQFEFQQYYQ
                     VTLYGREIEQEEQRPVGQIYMRCQYCLKSLDLLEKLDICCSNQPFHKVRDHWKGRCRH
                     CKAIE"
     gene            643. .945
                     /gene="E7"
     CDS             643. .945
                     /gene="E7"
                     /note="putative"
                     /codon_start=1
                     /product="transforming protein E7"
                     /db_xref="PID:g1020180"
                     /translation="MIGKQATLCDIVLEELVLPIDLHCHEELPELPEELEESVVEEEP
                     EYTPYKIVVYCGGCDTKLKLYILATLSGIRDFQTSLLGPVKLLCPTCREEIRNGRR"
     gene            932. .2758
                     /gene="E1"
     CDS             932. .2758
                     /gene="E1"
                     /note="putative"
                     /codon_start=1
                     /product="replication protein E1"
                     /db_xref="PID:g1020181"
                     /translation="MDDDKGTDTTDAKEGCSGWFMLEAACSDDSDLDNSLEKLFEDGT
                     ESDVSDLINDDDTAAQGNSRELLCQQQSEECEQQIQYLKRKYFSPKAVQQLSPRLQSM
                     NISPGHKSKRRLFVEHDSGLECSLNEAEDLTEEVEVPASAPAPAAQGGVGSGHYTSLL
                     RCNNVKAVLLGKFKDAFGVSYNELTRQFRSNKTCCKHWVLAIYAAKDELIDASKQLLQ
                     QHCTYLWLQTFSPMSLYLCCFNVGKSRETVMRLLSSMLQVNENHILSEPPKIRSMIAA
                     LFWYKGSMNPNVYAFGEYPEWIMTQTMIHHQTADSVQFDLSEMIQWAYDQDYVDECTI
                     AYQYARLADSNSNARAFLAHNSQAKYVRECAQMVRYYKRGEMRDMSISAWIHHCISKI
                     EGDGHWQDIVKFLRYQGLNFIVFLDKFRTFLKNFPKKNCLLICGPPDTGKSMFSMSLM
                     KALRGQVVSFANSKSHFWLQPLADAKLALLDDATEVCWQYIDAFLRNGLDGNMVSLDM
                     KHRAPCQMKFPPLIITSNISLKKEKKFPYLHSRIYEFEFPNKFPFDANDTPLFKLTDQ
                     SWASFFKRLWTQLELSDQEEEGENGETQRTFQCTTREVNGLI"
     gene            2700. .4010
                     /gene="E2"
     CDS             2700. .4010
                     /gene="E2"
                     /note="putative"
                     /codon_start=1
                     /product="regulatory protein E2"
                     /db_xref="PID:g1020182"
                     /translation="MEKLSERFSALQEKLMDLYESGVEDLETQIQHWKLLRQEQVLFY
                     YARRHGILRLGYQPVPTLATSESKAKDAIAMGLLLESLQKSQYAEEPWTLVETSLETV
                     KSPPADCFKKGPKSVEVYFDGDPENVMSYTVWSYIYYQTDDESWEKVEGHVDYTGAYY
                     IEGTFKTYYIKFETDAKRYGTTGHWEVHVNKDTVFTPVTSSTPPVGVASQNSAPEPAS
```

```
                     TSDSPQRSSQVTHRYGRKASSPTITTIRRQKRRERQRQETPTRRRKTRSRSRSTEQRG
                     GRATRRSLSRESAESPRRGGRGGGGPLTRSRSRSRSRTRESVDGGGVAPDEVGATLRS
                     IGRQHSGRLAQLLDAAKDPPVILLRGAANTLKCYRYRFRKKHAGSFQFISTTWSWVGG
                     HTTDRIGRSRILISFHTDREREKCLQQMKLPLGVEWSYGQFDDL"
     gene            3184. .3765
                     /gene="E4"
     CDS             <3184. .3765
                     /gene="E4"
                     /note="putative"
                     /codon_start=1
                     /db_xref="PID:g1020183"
                     /translation="KGPLKPIILNLKQMLNDMVQQDIGRCMLIKILCLPLLPVLRRQL
                     ESPPRTPHPNRHPPPTPHNGHHKSPTDTAEKHLVLQSPPSGGKKGERDKDKKPQQGEE
                     KPDQGPEAPSSGEGGPPDDPSPENPQNPPGGEGEVEGAPSPGPAQGRDPVHESLLTGV
                     ASRLTKWEQHFDQLVDSIVGDLRNYWTQLKTPQ"
     gene            4077. .5651
                     /gene="L2"
     CDS             4077. .5651
                     /gene="L2"
                     /note="putative"
                     /codon_start=1
                     /product="minor capsid protein L2"
                     /db_xref="PID:g1020184"
                     /translation="MARARRTKRASVTDIYKGCKASGTCPPDVINKVEQNTLADKILK
                     YGSVGVFFGGLGISTGKGTGGPTGYIPLGQGPGVRVGATPTVVRPGVIPEIIGPTELI
                     PVDSVTPIDPAAPSIVTLTDSSAGADLLPGEVETIAEVHPVPIDNVELDTPLVSGDRH
                     AILEVTDANPPFRRTVTRTQYHNPAFEIISESTPLIGESTPSDHVFVFEGSGGVQVGD
                     ANESIELDTFPSRYSFDIEEPTPPRRVSTPIERISQEFRTLRRALYNRRLTEQVQVRD
                     PLFIRSPSRLVRFQFDNPVFDEEVTQIFERDVAAVEEPPDRDFLDIERLGRPILTETA
                     EGRVRVSRLGQRASLSTRSGARVGARVHFFTDISTINAEEPIELELLGEHSGDSSVVQ
                     EPFESTILDVNIDNIPESLDTNIAETSVDYDSADLLLDNGVEDFSRSQLVIGPSDRSL
                     PSITVPQFESPRETIVYIQDIEGNTVVYPKYEERPTIILPTPSGPAIIQSPTHSSFDY
                     YLHPSLRRKKRKRKYL"
     gene            5662. .7194
                     /gene="L1"
     CDS             5662. .7194
                     /gene="L1"
                     /note="putative"
                     /codon_start=1
                     /product="major capsid protein L1"
                     /db_xref="PID:g1020185"
                     /translation="MTLWLPTSGKIYLPPTPPVARVQNTDEYVERTDIYYHAISDRLL
                     TVGHPYFDVRSSDGAKIEVPKVSGNQFRAFRVTFPDPNKFALGDMTIHDPERYRLVWA
                     CKGLEIGRGQPLGVGTTGHPLFNKLHDTENPTERQEGTSDDRRNVSFDPKQVQMFIIG
                     CIPCLGEYWDKAPVCEDAGSQVGLCPPLELKNGVIEDGDMFDIGFGNINNKTLSFNRS
                     DVSLDIVNEICKYPDFLTMSNDVYGDSCFFCARREQCYARHNFVRGGLVGDAIPDDAV
                     QQDHKYYLPAASQTALENSTYFPTVSGSLVTSDAQLFNRPFWLKRAQGHNNGILWNNQ
                     MFVTVADNTRNTNFSISVASDGTTVNYDAKKIREFMRHVEEYQLSFILQLCRIPLEAE
                     VLTQINAMNHGILENWQLGFVPTPDNSVHDTYRYLQSKATKCPDAVPDTQKEDPFGQY
                     TFWNVDMSEKLSLDLDQYPLGRKFLFQSGLQRARASARVSVKRSATRKTSKTVKRRKL
                     TS"
BASE COUNT     2315 a   1352 c   1614 g   2087 t
ORIGIN

  Query Match           47.6%;  Score 20;  DB 33;  Length 7368;
  Best Local Similarity  76.3%;  Pred. No. 8.14e+00;
  Matches    29;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    727 CCTGAACTTCCAGAAGAGTTAGAAGAATCAGTGGTAGA 764
          |||  || | ||||||| |||  |||| ||| ||||||
Cp    363 CCTTTACCTTCAGAAGATTTACCAGAACCAGAGGTAGA 326


RESULT   15
LOCUS       AE001306    13316 bp    DNA             BCT       02-SEP-1998
DEFINITION  Chlamydia trachomatis section 33 of 87 of the complete genome.
ACCESSION   AE001306 AE001273
NID         g3328748
KEYWORDS    .
SOURCE      Chlamydia trachomatis.
  ORGANISM  Chlamydia trachomatis
            Eubacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE   1  (bases 1 to 13316)
  AUTHORS   Stephens,R.S., Kalman,S., Lammel,C.J., Fan,J., Marathe,R.,
            Aravind,L., Mitchell,W.P., Olinger,L., Tatusov,R.L., Zhao,Q.,
            Koonin,E.V. and Davis,R.W.
  TITLE     Genome Sequence of an Obligate Intracellular Pathogen of Humans:
            Chlamydia trachomatis
  JOURNAL   Science (1998) In press
REFERENCE   2  (bases 1 to 13316)
  AUTHORS   Stephens,R.S., Kalman,S., Lammel,C.J., Fan,J., Marathe,R.,
            Aravind,L., Mitchell,W.P., Olinger,L., Tatusov,R.L., Zhao,Q.,
            Koonin,E.V. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAY-1998) Program in Infectious Diseases, University
            of California, 235 Warren Hall, Berkeley, CA 94720-7360, USA
FEATURES             Location/Qualifiers
     source          1. .13316
                     /organism="Chlamydia trachomatis"
                     /strain="D/UW-3/CX"
                     /db_xref="taxon:813"
     gene            136. .408
                     /gene="CT330"
     CDS             136. .408
                     /gene="CT330"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /db_xref="PID:g3328749"
                     /translation="MQEISVPEASHLDTVIQRLLVLQKALEDCHSVEAILAIYEEMFS
                     LIQEGETNLNHVEQLSFQLQLNPDGSVAVDTSGNPIKHPFIPGESV"
     gene            405. .2327
                     /note="CT331"
                     /gene="dxs"
     CDS             405. .2327
                     /gene="dxs"
                     /codon_start=1
                     /transl_table=11
                     /product="Transketolase"
                     /db_xref="PID:g3328750"
                     /translation="MTYSLLPHIHSPQDLHALSLDKLPVLCDEIRNKIIESLSLTGGH
                     LASNLGGVELTVALHYVFSSPDDQFIFDVGHQSYVHKLLTGRNTEAFSNIRHDNGLSG
                     FTTPQESNHDIFFSGHAGNALSLALGLAKGSSNSSSHILPILGDAAFSCGLTLEALNN
                     IPADLSKFIIVLNDNQMSISENVGNIPQGISHWMYPQKISKLSQKIHSWIQNLPSFLH
                     KKKTLSHKVDIALKSLSHPLFEQFGLHYVGPIDGHNVKKLVQALQMIKDQPQPILFHV
                     CTVKGNGLAEAERDPIRYHGVKAHFQNTSLKKTSGNVELQTPISFPQHAGNILCRLGK
                     KYPQLQVVTPAMSLGSCLEDFRKQFPDRFTDVGIAEGHAVTFSAGIARSGTPVCCSIY
                     STFLHRAMDNVFHDVCMQELPVIFAIDRAGLAFHDGRSHHGIYDLGFLCSMPNMVICQ
                     PRNALVLERLFFSSLLWKSPCAIRYPNIPANEKASNSFFPFSPILPGEAEILCQGDDL
                     LLIALGHMCNTALTVKEHLLDYGISTTVVDPIFIKPLDRKLLQSLLTHHSKVIILEEH
                     SIHGGLGSEFLLFLNQHNIKADVLSLGVPDMFIPHGNPETILNLIGLTSDHITQRILS
                     HFKFSTPIPIERFFKA"
     gene            2424. .3881
                     /note="CT332"
                     /gene="pykF"
     CDS             2424. .3881
                     /gene="pykF"
                     /codon_start=1
                     /transl_table=11
                     /product="Pyruvate Kinase"
                     /db_xref="PID:g3328751"
                     /translation="MIARTKIICTIGPATNTPEMLEKLLDAGMNVARLNFSHGTHESH
                     GRTIAILKELREKRQVPLAIMLDTKGPEIRLGQVESPIKVQPGDRLTLVSKEILGSKE
                     SGVTLYPSCVFPYVRERAPVLIDDGYIQAVVVNAQEHMVEIEFQNSGEIKSNKSLSIK
                     DIDVALPFMTEKDIADLKFGVEQELDLIAASFVRCNEDIDSMRKVLESFGRPNMPIIA
                     KIENHLGVQNFQEIARAADGIMIARGDLGIELSIVEVPGLQKFMARASRETGRFCITA
                     TQMLESMIRNPLPTRAEVSDVANAIYDGTSAVMLSGETALGAHPVHAVKTMRSIIQET
                     EKTFDYHAFFQLNDKNSALKVSPYLEAIGFSGIQIAEKASAKAIIVYTQTGGSPMFLS
                     KYRPYLPIIAVTPNRNVYYRLAVEWGVYPMLTLESNRTVWRHQACVYGVEKGILSNYD
                     KILVFSRGAGMQDTNNLTLTTVHDVLSPSLDEIVP"
     gene            3904. .9264
                     /note="CT333"
                     /gene="uvrA"
     CDS             3904. .9264
                     /gene="uvrA"
                     /codon_start=1
```

```
                     /transl_table=11
                     /product="Excinuclease ABC Subunit A"
                     /db_xref="PID:g3328752"
                     /translation="MSSIVRLSGITVRNLKNITVEFCPREIVLFTGVSGSGKSSLAFN
                     TIYAAGRKRYLATLPSFFATKLDSLPDPSVKKIEGLSPTVAVKQNFFEQHIHATVGST
                     TEINSYLALLFSLEGQAYDPVTLHPLTLYSKEKILSEIAAIPDGTQLTLLAPLPARDV
                     LTVRECLRQGFTKVLIDEEVVPIHKFLTTGIPVPGQLIVDTLIKNTSNTPRLKVSLFT
                     TLDIGHGECSLHFDNQKRVFSTQATLPETQQTYAPLSPDLFSSHSHKDRCPQCQGSGI
                     FISIDDPCIIQQTLSIEDNCCPFAGNCSTFLYRTIYQSLADSLGFSLSTPWQNLSPEI
                     QHIFLYGKEGLSIPVKLFDGTLGKKTQTHKQWKGVLNEIGEKIRFSNKPARYLPKGTS
                     YTVCPRCNKTGLSDYANATRWHGKSFADFQQMSLQELFIFLNQLPEKELAIEEVIQGF
                     KSRLTILIDLGLPYLSPERAIDTLSGGERERTALAKHLGAELIGVMYILDEPSIGLHP
                     QDTHKLMNVIRRLRDQGNTVLLVEHDEQMISLADRVIDIGPGAGIFGGEVVFNGSPRE
                     FLAKSDSLTAQYLRQEQHISVPAKRTNSLGTITLSKANKHNLKDLTVSIPLGQMTVVT
                     GVSGSGKSSLINDTLVPCVEEFIEQGFCSTLSIQGALSRLVHINRDLPGRSQRSIPLT
                     YIKAFDEVRQLFAEQPRCKTLGLTKSHFSFNTPLGACAECGGLGSITTTDNRDSITCP
                     SCLGKRFLPQVLEVRYKNKTIADILEMTAYEAKNFFLDEPSIHQKIETLCTLGLQYLP
                     LGRPLYSLSGGEIQRLKLAYELLAPVKHPTLYVLDEPTTGLHTHDVKQLIYVVQSLIH
                     QGHSVIIIEHNMHVVKVADYIILELGPEGGNKGGYLIASCSPEELIHKHTPTAIALRPF
                     LSSPQELPYLPDPSPKPPVPAAITIANAHQHNLKHIDLSIPRYALTAVTGPSASGKHS
                     LVFDILHAAGNIAYAELFPPYIRQALIKKTPLPAVDKVTGLSPVIAIEKTSASRNSNH
                     SVASALEISEMLESLFTRIGHPYSPISGDTLRTISPETIAEELLTHYTKGYVTITVPF
                     PKEEEFFSYTQEMLQEGFLKLYANEQFYDLDGPFPTSLENPALVIHHVKILEKNLPSL
                     LASLTLAFSKASSVCLHIEYAGTSLSKTYRQGLQDASGNLFPNIETPSILNHESYLCP
                     LCHGKGFLSTCSILPHKKRFAQHTPISLFTSLFPNQDPSPIYPLLNELGIPSIALFQE
                     IDILSFESLCLGTQQHLGLNALCTKAMLMESEEDFPPDLISKTPCNQCQGLGVYTYKH
                     CTRIHSISLSEIYQSNVAFLKKLLLSVEEEPSLVQDILSRLDLLDRVGLGYVILGQEQ
                     QSLSDGEHYRLLLAKAFSSGLTDVIYLLEDPLAGIHPEDAPCLLAVIKKLVTNHNTVI
                     VTDREGSLAEHADYLLHLGPEPGPNGGYLLSTSALKQSQPVLCNTRSSEETPQLSVSV
                     STSTIQIENLAFPLQRLSTISGVSGSGKTTLLNSIYEHGCAILEKDPSVFSEIIFLDS
                     YPQIASSRSDISTYFDIAPSLRNFYASLTQAKALNISVSMFSPNTKQGQCSDCWGLGY
                     QWIDRAFYAMEKRPCPTCGGFRVQPLIQEVVYEGKHFGQLLQASLNEVAETFSFLKKI
                     QKPLHTLITNDLGYLSLGQNMASLSLSEKIAIKLTKHLFLPPKHPTLFLLDEIATSLD
                     NQQQSALLVQLNTLVSLGHTVVIIENHPAFSQQADFWIQMGRKTDNHILFAGPNPNLS
                     FT"
gene                 9482. .10882
                     /note="CT334"
                     /gene="dnaX_2"
CDS                  9482. .10882
                     /gene="dnaX_2"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA Pol III Gamma and Tau"
                     /db_xref="PID:g3328753"
                     /translation="MTSYQVSSRRYRPNNFSEILGQDIVVSILKNSLRLNRAAHAYIF
                     SGIRGTGKTTLARVFAKALNCQNPTQDQEPCNQCAICKEISLGTSMDVIEIDGASHRG
                     IEDIRQINETVLFVPSKSRYKIYIIDEVHMLTKEAFNSLLKTLEEPPVHVKFFLATTE
                     IAKIPNTISSRCQKMLLKRISEETIIDKLATIAKQEGTETSREALLPIAKAAQGSLRD
                     AESLYDYVVGLFPESLDPDSTAKALGILSEDSLYQLAEAITTQNYEQALTPVSDAMQM
                     GVAPAHFLQDVTLLFRSFLLKQYSKQFHSVATKYSSEILLEILDFLGESARHINLALF
                     EKTFLETVIIRLLRIYSRPTFSQLVSQIRQPAPTQIRPIPTASELESCPPQIKSSNTE
                     ETKSSPQPTAVAKLSQGSLLTTASSQTKKETIHQETLPPSLDKISPSESAAIDTLLQF
                     AVVEFSGVLTKEPKHG"
gene                 10875. .11165
                     /gene="CT335"
CDS                  10875. .11165
                     /gene="CT335"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /db_xref="PID:g3328754"
                     /translation="MGSGYAKKKKEAKLMERQFMEMEASLEQKRFSGEAGNGLVSVTI
                     NGKCDLVDVRIKPDCLDPEDPEVVADLFRAAFKAAKAALDSEMSAMQMGMPF"
gene                 complement(11175. .12890)
                     /note="CT336"
                     /gene="ptsI"
CDS                  complement(11175. .12890)
                     /gene="ptsI"
                     /codon_start=1
                     /transl_table=11
                     /product="PTS PEP Phosphotransferase"
                     /db_xref="PID:g3328755"
                     /translation="MSVTGQDNKELQQEFVIVGEPIVPGIGLGKALLLGKSSLRIREL
                     TLPQEEVEHEISRYYKALKRSRSDLAALEKEAKGKQGYQEIASILQAHLEIIKDPLLT
                     EEVVKTIRKDRKNAEFVFSSVMGEIEKSLCAVQKTTATRVDRVQDIHDISNRVIGHLC
                     CQHKSSLGEFDQNLIVFSEELTPSEAANANPEYIRGFVSLEGAKTSHTAIVSLAKNIP
                     YVANFTTELWDTIKEFSGTLVLINGDKGEITFNPQLSTIQTYYRKQASVSVTVPVQVQ
                     TGKNLPLISLSAQIVSTEELPMIERESPGTSVGLFRSEFMAFSLGRLPCVEEQADQYA
                     QLVQFQCSDIHVLRLFDFGEDKECPCISSSHRSVRWLLEQEKVLKEQLQAIAIVSRIG
                     RLKVLIPGVIDASEIALVKRLFQEEIRLLKGISENILWGSMIEIPSAVWMIEEILQES
                     SFVALGTNDLAQYTLGTSRERSLLGERSRVPHPSVIRMIHHVVEQAKQKNVPVSVCGE
                     MAGDPALLPMFLGLGVKELSAVIPAINSLKMRLLDLNSRECSRLTKQLLRAKTYEEVH
                     QLLYV"
gene                 complement(12890. .13219)
                     /note="CT337"
                     /gene="ptsH"
CDS                  complement(12890. .13219)
                     /gene="ptsH"
                     /codon_start=1
                     /transl_table=11
                     /product="PTS Phosphocarrier Protein Hpr"
                     /db_xref="PID:g3328756"
                     /translation="MSGGNDSVLHRPVEAPLMEDGELSAVFTIRNSSGIHVRPAGTIV
                     KLFEGEECEATLTYLGKTVNARSVMSILMLGASYNGEVAVHIKGPSASRVMQKLSEVF
                     NSGFGEL"
BASE COUNT     3888 a   3167 c   2371 g   3890 t
ORIGIN

  Query Match           47.6%;  Score 20;  DB 16;  Length 13316;
  Best Local Similarity  95.5%;  Pred. No. 8.14e+00;
  Matches    21;  Conservative    0;  Mismatches   1;  Indels   0;  Gaps   0;

Db    5781 CTCTGGTTCTGGGAAATCTTCT 5802
           ||||||||||||| |||||||||
Qy     331 CTCTGGTTCTGGTAAATCTTCT 352


Search completed: Sat Jan  9 17:12:39 1999
Job time : 107 secs.
```

****

FASTA (TM)

****

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Jan 9 17:20:30 1999; MasPar time 20.09 Seconds
284.402 Million cell updates/sec

Tabular output not generated.

Title: >US-08-704-178-1
Description: (322-363) from US08704178.seq (4 of 4)
Perfect Score: 42
N.A. Sequence: 322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363
Comp: TCCAAGATGGAGACCAAGACCATTTAGAAGACTTCCATTTCC

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 5.992; Variance 4.139; scale 1.448

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
	1	42	100.0	711	9	Q55180	Sequence encoding the	4.08e-12
	2	42	100.0	711	18	T17728	Anti-erbB2 scFv cDNA.	4.08e-12
	3	42	100.0	711	29	T65006	Single-chain anti-erb	4.08e-12
	4	42	100.0	761	7	Q43289	Sequence encoding a 4	4.08e-12
	5	42	100.0	803	3	Q14831	Qm212 single chain an	4.08e-12
	6	38	90.5	720	29	T65007	Single-chain anti-erb	6.29e-10
	7	38	90.5	720	9	Q55181	Sequence encoding the	6.29e-10
	8	38	90.5	731	7	Q43287	Sequence encoding a m	6.29e-10
	9	38	90.5	770	7	Q43290	Sequence encoding a 4	6.29e-10
c	10	32	76.2	91	9	Q51746	Oligonucleotide probe	1.03e-06
	11	28	66.7	744	7	Q43288	Sequence encoding a m	1.24e-04
	12	25	59.5	91	9	Q51746	Oligonucleotide probe	4.10e-03
	13	24	57.1	797	22	T29058	CC49/212 SCA PLAP DNA	1.29e-02
	14	24	57.1	803	22	T29059	PLAP CC49/212 SCA DNA	1.29e-02
	15	24	57.1	1460	7	Q43291	Sequence encoding div	1.29e-02
	16	23	54.8	738	7	Q43293	Sequence encoding sin	4.00e-02
	17	23	54.8	738	11	Q66841	CC49 VL / 217 / 4-4-2	4.00e-02
	18	21	50.0	39	7	Q51787	Mixed oligonucleotide	3.70e-01
	19	21	50.0	47	35	T86394	Sp6 anti-TNP antibody	3.70e-01
	20	21	50.0	66	8	Q49187	scFvR construction pr	3.70e-01
	21	20	47.6	36	2	Q11195	Ballast Constituent c	1.10e+00
c	22	20	47.6	50	13	Q80512	Primer used in immuno	1.10e+00
	23	20	47.6	725	11	Q66840	4-4-20 VL/ 217 / CC49	1.10e+00
	24	20	47.6	725	7	Q43292	Sequence encoding sin	1.10e+00
	25	20	47.6	5600	3	Q22963	Sequence of pHCV-31 e	1.10e+00
	26	20	47.6	5600	11	Q62663	pHCV-31 plasmid expre	1.10e+00
	27	20	47.6	5600	6	Q38266	HCV CKS-33-BCD fusion	1.10e+00
	28	20	47.6	5600	6	Q38251	HCV CKS-33-BCD fusion	1.10e+00
	29	20	47.6	5600	3	Q21678	pHCV-31 - recombinant	1.10e+00
	30	20	47.6	5600	6	Q38236	HCV CKS-33-BCD expres	1.10e+00
	31	19	45.2	5362	39	T97296	Mouse additional sex	3.21e+00
	32	19	45.2	5362	39	T97294	Mouse additional sex	3.21e+00
c	33	19	45.2	9202	2	N70608	VISNA sheep lentiviru	3.21e+00
	34	18	42.9	33	7	Q46603	Mixed oligonucleotide	9.18e+00
c	35	18	42.9	1144	10	Q44800	Encodes single-strand	9.18e+00
c	36	18	42.9	1144	38	T99264	Human PUR-alpha gene.	9.18e+00
	37	18	42.9	2956	37	T99140	Full length cytochrom	9.18e+00
c	38	18	42.9	3279	9	Q50946	Sequence encoding pro	9.18e+00
c	39	18	42.9	3279	9	Q51556	Sequence encoding pro	9.18e+00
	40	18	42.9	3388	34	T84543	Human steroid recepto	9.18e+00
c	41	18	42.9	3399	20	T05868	Chicken leucocytozoan	9.18e+00
	42	18	42.9	134525	1	Q04525	Total base sequence o	9.18e+00
	43	17	40.5	563	1	N90964	Sequence in a recombi	2.57e+01
	44	17	40.5	668	1	Q06496	IFN-alpha 61 gene and	2.57e+01
	45	17	40.5	3537	1	Q06831	Sequence encoding B.t	2.57e+01

ALIGNMENTS

RESULT 1
ID Q55180 standard; cDNA; 711 BP.
AC Q55180;
DT 21-JUL-1994 (first entry)
DE Sequence encoding the single chain anti-erbB2 antibody, Ab no.23.
KW Single chain anti-erbB1 antibody; cancer therapy; prevention;
KW monoclonal antibody; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT cds 1..171
FT /*tag= a
PN WO9400136-A.
PD 06-JAN-1994.
PF 21-OCT-1992; U08545.
PR 30-JUN-1992; US-906555.
PA (MOLE-) MOLECULAR ONCOLOGY INC.
PI Kasprzyk PG, King CR;
DR WPI; 94-025878/03.
DR P-PSDB; R45442.
PT Treatment of malignancies over-expressing ERB-[2 - using at least
PT 2 monoclonal antibodies which recognise different epitopes on
PT gp185
PS Example; Fig 7; 37pp; English.
CC The source of human erbB-2 protein for the prodn. of antibodies no.
CC 23 and 21 (Q55180/R45442; Q55181/R45443) is a NIH/3T3 cell
CC engineered to express the human erbB-2 protein on its surface (N/
CC erbB-2). Abs no. 21 and 23 are directed against the extracellular
CC domain of gp185 erbB-2. Nude mice manipulated to produce rapidly
CC growing tumours were used in a trial of the efficacy of the Abs. In
CC animals given a combination of the 2 Abs, tumours completely
CC regressed after 11 days.
SQ Sequence 711 BP; 175 A; 183 C; 187 G; 165 T;

Query Match 100.0%; Score 42; DB 9; Length 711;
Best Local Similarity 100.0%; Pred. No. 4.08e-12;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db    321 aggttctacctctggttctggtaaatcttctgaaggtaaagg 362
          ||||||||||||||||||||||||||||||||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT    2
ID    T17728 standard; cDNA; 711 BP.
AC    T17728;
DT    21-MAY-1996  (first entry)
DE    Anti-erbB2 scFv cDNA.
KW    Oncoprotein; erbB2; cell proliferation; tumour; cancer;
KW    intracellular antibody homologue; single chain antibody; scFv;
KW    gene therapy; ds.
OS    Synthetic.
PN    WO9607321-A1.
PD    14-MAR-1996.
PF    23-AUG-1995; U10740.
PR    06-SEP-1994; US-301339.
PR    06-JUN-1995; US-468252.
PA    (UABR-) UAB RES FOUND.
PI    Curiel DT,  Deshane J;
DR    WPI; 96-171307/17.
DR    P-PSDB; R94020.
PT    Inhibition of proliferation or survival of, esp. malignant erbB2,
PT    cells - by introducing nucleic acid mol. encoding antibody homologue
PT    which is expressed and binds, pref. erbB2, protein intracellularly
PS    Claim 42; Page 29-30; 48pp; English.
CC    A nucleic acid comprises a first sequence encoding a signal peptide
CC    (R94019) linked to a second sequence (T17728) encoding a single
CC    chain Fv fragment (R94020) that binds a human erbB2 oncoprotein.
CC    The anti-erbB2 sFv portion is obtained by PCR using e23scFv
CC    plasmid as template.  The signal peptide directs the scFv to the
CC    endoplasmic reticulum.  The nucleic acid is incorporated into a
CC    plasmid or viral vector to facilitate expression of the scFv antibody
CC    homologue within e.g. an epithelial carcinoma cell.  Intracellular
CC    expression of the homologue inhibits surface expression of erbB2 and
CC    thereby inhibits cell proliferation and cell survival and decreases
CC    tumorigenicity.
SQ    Sequence   711 BP;    175 A;    183 C;    189 G;    164 T;

  Query Match          100.0%;  Score 42;  DB 18;  Length 711;
  Best Local Similarity 100.0%;  Pred. No. 4.08e-12;
  Matches    42;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    321 aggttctacctctggttctggtaaatcttctgaaggtaaagg 362
          ||||||||||||||||||||||||||||||||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT    3
ID    T65006 standard; cDNA; 711 BP.
AC    T65006;
DT    05-JUN-1997  (first entry)
DE    Single-chain anti-erbB2 antibody e23(Fv) cDNA.
KW    Single chain antibody; variable region; light chain; heavy chain;
KW    breast cancer; ovarian cancer; non-small cell lung carcinoma;
KW    immunodiagnosis; treatment; cytotoxic agent; erbB-2; ds.
OS    Mus musculus.
OS    Synthetic.
FH    Key             Location/Qualifiers
FT    mat_peptide     1..711
FT                    /*tag=  a
FT                    /product= e23(Fv)
PN    US5587458-A.
PD    24-DEC-1996.
PF    07-OCT-1991; 772270.
PR    07-OCT-1991; US-772270.
PR    30-JUN-1992; US-906555.
PR    14-MAY-1993; US-061092.
PA    (ARON-) ARONEX PHARM INC.
PI    Bird RE,  Kasprzyk PG,  King CR;
DR    WPI; 97-064831/06.
DR    P-PSDB; W15185.
PT    Single chain antibodies specific for erbB-2 protein, gp185  - with
PT    labels or cytotoxin, useful for detection and treatment of tumour
PT    cells expressing this protein
PS    Example 8; Columns 25-28; 28pp; English.
CC    The present cDNA sequence codes for a claimed single-chain antibody,
CC    designated e23(Fv), which binds to erbB-2. Monoclonal antibody e23
CC    was generated by immunising mice with N/erbB-2 cells overexpressing
CC    the gp185 protein, removing spleen cells and producing hybridomas
CC    by standard techniques. Messenger RNA coding for the anti-erbB-2
CC    monoclonal antibody was isolated and converted to cDNA. Regions
CC    coding for the heavy- and light- chain variable regions were then
CC    amplified by PCR and joined via a sequence encoding a peptide
CC    linker. The resulting single-chain antibody is useful for in vitro
CC    diagnosis of tumour cells which overexpress the erbB-2 gp185
CC    marker, e.g. breast, ovarian and non-small cell lung carcinomas,
CC    and, when coupled to a cytotoxic agent, to treat such tumours.
SQ    Sequence   711 BP;    175 A;    182 C;    189 G;    164 T;

  Query Match          100.0%;  Score 42;  DB 29;  Length 711;
  Best Local Similarity 100.0%;  Pred. No. 4.08e-12;
  Matches    42;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    321 aggttctacctctggttctggtaaatcttctgaaggtaaagg 362
          ||||||||||||||||||||||||||||||||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT    4
ID    Q43289 standard; DNA; 761 BP.
AC    Q43289;
DT    27-SEP-1993  (first entry)
DE    Sequence encoding a 4-4-20/212 single-chain antigen-binding
DE    protein with a single cysteine hinge.
KW    Antibody; multivalent; variable region; heavy chain; light chain;
KW    linker; ss.
OS    Synthetic.
FH    Key             Location/Qualifiers
FT    cds             1..761
FT                    /*tag=  a
FT    cds             1..336
FT                    /*tag=  b
FT                    /product= 4-4-20 VL
FT    cds             337..378
FT                    /*tag=  c
FT                    /product= 212 linker
FT    cds             379..729
FT                    /*tag=  d
FT                    /product= CC49 VH
FT    cds             730..750
FT                    /*tag=  e
FT                    /product= hinge
PN    WO9311161-A.
PD    10-JUN-1993.
PF    20-NOV-1992; U09965.
PR    25-NOV-1991; US-796936.
PA    (ENZO-) ENZON INC.
PI    Bird RE,  Filpula D,  Hardman K,  Rollence M,  Whitlow MD,
PI    Wood JF;
DR    WPI; 93-196999/24.
DR    P-PSDB; R37647.
PT    New multivalent antigen-binding protein e.g. contg. CC49 Mab
PT    regions - useful in diagnosis, for destroying blood clots and
PT    targetting cytotoxic agents or enzymes to tumour cells
PS    Example; Fig 15A; 118pp; English.
CC    Free cysteines were engineered into the C-terminal of the
CC    4-4-20/212 single-chain antigen-binding protein, in order to
CC    chemically crosslink the protein. The design was based on the
CC    hinge region found in antibodies between the CH1 and CH2 regions.
CC    In order to try to reduce antigenicity in humans, the hinge
CC    sequence of the most common IgG class, IgG1, was chosen. In the
CC    design for the hinge region, the C-terminal serine in the
CC    4-4-20/212 singl-chain antigen-binding protein was made the first
CC    serine of the hinge and the second residue of the hinge was changed
```

```
CC   from a cysteine to a serine. The hinge regions were added by
CC   introduction of a BstE II restriction site in the 3'-terminus
CC   of the gene encoding the 4-4-20/212 single-chain antigen-binding
CC   protein.
SQ   Sequence   761 BP;    212 A;    162 C;    187 G;    200 T;

  Query Match          100.0%;  Score 42;  DB 7;  Length 761;
  Best Local Similarity 100.0%;  Pred. No. 4.08e-12;
  Matches    42;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    336 aggttctacctctggttctggtaaatcttctgaaggtaaagg 377
          |||||||||||||||||||||||||||||||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT   5
ID   Q14831 standard; DNA; 803 BP.
AC   Q14831;
DT   17-FEB-1992  (first entry)
DE   Qm212 single chain antigen binding protein with metal binding site.
KW   Antigen binding; metal cations; ss.
FH   Key             Location/Qualifiers
FT   cds             1..272
FT                   /*tag=  a
PN   WO9116912-A.
PD   14-NOV-1991.
PF   07-MAY-1991; U03149.
PR   08-MAY-1990; US-521258.
PR   18-JUN-1990; US-539980.
PA   (SCRI-) SCRIPPS CLINIC & RE.
PI   Lerner RA, Roberts VN, Getzoff ED, Tainer JA, Benkovic SJ;
DR   WPI; 91-353518/48.
DR   P-PSDB;R15055.
PT   New metal binding proteins - comprising variable domain of
PT   immunoglobulin and 3 contact aminoacid residues as metal binding
PT   site
PS   Disclosure; Fig 1; 73pp; English.
CC   The sequence encodes Qm212, a single chain antigen binding protein
CC   having a metal binding site which is derived from the 4-4-20/212
CC   protein (Bird et al., Science, 242:423-426 {1988}). Qm212 differs
CC   from 4-4-20/212 by amino acid substitutions at residue positions
CC   60, 62, 115 and 117 to provide for contact amino acid residues,
CC   where three of these contact residues form a metal binding site.
CC   The organism form which the gene originated is not given in the
CC   specification. The protein may be used to selectively partition
CC   and/or remove metal cations from fluids. Preselected antigens can
CC   be used to remove metal-protein complexes from a liquid phase or as
CC   indicators of metal complex formation. The metal-protein complexes
CC   can also be used to promote a predetermined chemical reaction, e.g.
CC   hydrolysis of a peptide bond.
SQ   Sequence   803 BP;    224 A;    175 C;    197 G;    207 T;

  Query Match          100.0%;  Score 42;  DB 3;  Length 803;
  Best Local Similarity 100.0%;  Pred. No. 4.08e-12;
  Matches    42;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    399 aggttctacctctggttctggtaaatcttctgaaggtaaagg 440
          |||||||||||||||||||||||||||||||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT   6
ID   T65007 standard; cDNA; 720 BP.
AC   T65007;
DT   05-JUN-1997  (first entry)
DE   Single-chain anti-erbB2 antibody e21(Fv) cDNA.
KW   Single chain antibody; variable region; light chain; heavy chain;
KW   breast cancer; ovarian cancer; non-small cell lung carcinoma;
KW   immunodiagnosis; treatment; cytotoxic agent; erbB-2; ds.
OS   Mus musculus.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     1..720
FT                   /*tag=  a
FT                   /product= e21(Fv)
PN   US5587458-A.
PD   24-DEC-1996.
PF   07-OCT-1991; 772270.
PR   07-OCT-1991; US-772270.
PR   30-JUN-1992; US-906555.
PR   14-MAY-1993; US-061092.
PA   (ARON-) ARONEX PHARM INC.
PI   Bird RE,  Kasprzyk PG,  King CR;
DR   WPI; 97-064831/06.
DR   P-PSDB; W15186.
PT   Single chain antibodies specific for erbB-2 protein, gp185  - with
PT   labels or cytotoxin, useful for detection and treatment of tumour
PT   cells expressing this protein
PS   Example 9; Columns 27-30; 28pp; English.
CC   The present cDNA sequence codes for a claimed single-chain antibody,
CC   designated e21(Fv), which binds to erbB-2. Monoclonal antibody e21
CC   was generated by immunising mice with N/erbB-2 cells overexpressing
CC   the gp185 protein, removing spleen cells and producing hybridomas
CC   by standard techniques. Messenger RNA coding for the anti-erbB-2
CC   monoclonal antibody was isolated and converted to cDNA. Regions
CC   coding for the heavy- and light- chain variable regions were then
CC   amplified by PCR and joined via a sequence encoding a peptide
CC   linker. The resulting single-chain antibody is useful for in vitro
CC   diagnosis of tumour cells which overexpress the erbB-2 gp185
CC   marker, e.g. breast, ovarian and non-small cell lung carcinomas,
CC   and, when coupled to a cytotoxic agent, to treat such tumours.
SQ   Sequence   720 BP;    184 A;    178 C;    179 G;    179 T;

  Query Match          90.5%;  Score 38;  DB 29;  Length 720;
  Best Local Similarity 95.2%;  Pred. No. 6.29e-10;
  Matches    40;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Db    315 aggttctacctccggatctggtaaatcttctgaaggtaaagg 356
          |||||||||||| || ||||||||||||||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT   7
ID   Q55181 standard; cDNA; 720 BP.
AC   Q55181;
DT   21-JUL-1994  (first entry)
DE   Sequence encoding the single chain anti-erbB2 antibody, Ab no.21.
KW   Single chain anti-erbB1 antibody; cancer therapy; prevention;
KW   monoclonal antibody; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   cds             1..171
FT                   /*tag=  a
PN   WO9400136-A.
PD   06-JAN-1994.
PF   21-OCT-1992; U08545.
PR   30-JUN-1992; US-906555.
PA   (MOLE-) MOLECULAR ONCOLOGY INC.
PI   Kasprzyk PG,  King CR;
DR   WPI; 94-025878/03.
DR   P-PSDB; R45443.
PT   Treatment of malignancies over-expressing ERB-[2 - using at least
PT   2 monoclonal antibodies which recognise different epitopes on
PT   gp185
PS   Example; Fig 8; 37pp; English.
CC   The source of human erbB-2 protein for the prodn. of antibodies no.
CC   23 and 21 (Q55180/R45442; Q55181/R45443) is a NIH/3T3 cell
CC   engineered to express the human erbB-2 protein on its surface (N/
CC   erbB-2). Abs no. 21 and 23 are directed against the extracellular
CC   domain of gp185 erbB-2. Nude mice manipulated to produce rapidly
CC   growing tumours were used in a trial of the efficacy of the Abs. In
CC   animals given a combination of the 2 Abs, tumours completely
CC   regressed after 11 days.
SQ   Sequence   720 BP;    184 A;    178 C;    179 G;    179 T;

  Query Match          90.5%;  Score 38;  DB 9;  Length 720;
```

```
  Best Local Similarity  95.2%;  Pred. No. 6.29e-10;
  Matches    40;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Db    315 aggttctacctccggatctggtaaatcttctgaaggtaaagg 356
          ||||||||||||  || ||||||||||||||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT    8
ID   Q43287 standard; DNA; 731 BP.
AC   Q43287;
DT   27-SEP-1993  (first entry)
DE   Sequence encoding a multivalent antigen-binding single-chain protein
DE   comprising the 4-4-20 VL region connected through the 212 linker
DE   polypeptide to the CC49 VH region.
KW   Antibody; multivalent; variable region; heavy chain; light chain;
KW   linker; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   cds             1..729
FT                   /*tag=  a
FT   cds             1..336
FT                   /*tag=  b
FT                   /product= 4-4-20 VL
FT   cds             337..378
FT                   /*tag=  c
FT                   /product= 212 linker
FT   cds             379..723
FT                   /*tag=  d
FT                   /product= CC49 VH
PN   WO9311161-A.
PD   10-JUN-1993.
PF   20-NOV-1992; U09965.
PR   25-NOV-1991; US-796936.
PA   (ENZO-) ENZON INC.
PI   Bird RE,  Filpula D,  Hardman K,  Rollence M,  Whitlow MD,
PI   Wood JF;
DR   WPI; 93-196999/24.
DR   P-PSDB; R37645.
PT   New multivalent antigen-binding protein e.g. contg. CC49 Mab
PT   regions - useful in diagnosis, for destroying blood clots and
PT   targetting cytotoxic agents or enzymes to tumour cells
PS   Example; Fig 10A; 118pp; English.
CC   "4-4-20 VL" means the variable region of the light chain of the
CC   4-4-20 mouse monoclonal antibody (Bird, R.E. et al., Science 242:423
CC   (1988)). The number "212" refers to a specific 14-residue
CC   polypeptide linker that links the 4-4-20 VL and the CC49 VH. See
CC   Bedzyk, W.D. et al., J. Biol. Chem. 265:18615-18620 (1990). "CC49 VH"
CC   is the variable region of the heavy chain of the CC49 antibody,
CC   which binds to the TAG-72 antigen.
SQ   Sequence   731 BP;    190 A;    176 C;    180 G;    185 T;

  Query Match           90.5%;  Score 38;  DB 7;  Length 731;
  Best Local Similarity  95.2%;  Pred. No. 6.29e-10;
  Matches    40;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Db    336 aggttctacctctggttctggtaaatcctctgaaggcaaagg 377
          |||||||||||||||||||||||||| |||||||| |||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT    9
ID   Q43290 standard; DNA; 770 BP.
AC   Q43290;
DT   27-SEP-1993  (first entry)
DE   Sequence encoding a 4-4-20/212 single-chain antigen-binding
DE   protein with the two cysteine hinge.
KW   Antibody; multivalent; variable region; heavy chain; light chain;
KW   linker; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   cds             1..765
FT                   /*tag=  a
FT   cds             1..336
FT                   /*tag=  b
FT                   /product= 4-4-20 VL
FT   cds             337..378
FT                   /*tag=  c
FT                   /product= 212 linker
FT   cds             379..729
FT                   /*tag=  d
FT                   /product= CC49 VH
FT   cds             730..759
FT                   /*tag=  e
FT                   /product= hinge
PN   WO9311161-A.
PD   10-JUN-1993.
PF   20-NOV-1992; U09965.
PR   25-NOV-1991; US-796936.
PA   (ENZO-) ENZON INC.
PI   Bird RE,  Filpula D,  Hardman K,  Rollence M,  Whitlow MD,
PI   Wood JF;
DR   WPI; 93-196999/24.
DR   P-PSDB; R37648.
PT   New multivalent antigen-binding protein e.g. contg. CC49 Mab
PT   regions - useful in diagnosis, for destroying blood clots and
PT   targetting cytotoxic agents or enzymes to tumour cells
PS   Example; Fig 15B; 118pp; English.
CC   Free cysteines were engineered into the C-terminal of the
CC   4-4-20/212 single-chain antigen-binding protein, in order to
CC   chemically crosslink the protein. The design was based on the
CC   hinge region found in antibodies between the CH1 and CH2 regions.
CC   In order to try to reduce antigenicity in humans, the hinge
CC   sequence of the most common IgG class, IgG1, was chosen. In the
CC   design for the hinge region, the C-terminal serine in the
CC   4-4-20/212 singl-chain antigen-binding protein was made the first
CC   serine of the hinge and the second residue of the hinge was changed
CC   from a cysteine to a serine. The hinge regions were added by
CC   introduction of a BstE II restriction site in the 3'-terminus
CC   of the gene encoding the 4-4-20/212 single-chain antigen-binding
CC   protein.
SQ   Sequence   770 BP;    213 A;    166 C;    188 G;    203 T;

  Query Match           90.5%;  Score 38;  DB 7;  Length 770;
  Best Local Similarity  95.2%;  Pred. No. 6.29e-10;
  Matches    40;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Db    336 aggttctacgtgtggttctggtaaatcttctgaaggtaaagg 377
          ||||||||| | ||||||||||||||||||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT   10
ID   Q51746 standard; cDNA; 91 BP.
AC   Q51746;
DT   31-MAY-1994  (first entry)
DE   Oligonucleotide probe MK14-A
KW   Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW   ss.
OS   Synthetic.
PN   EP-571911-A.
PD   01-DEC-1993.
PF   24-MAY-1993; 108325.
PR   26-MAY-1992; US-889651.
PA   (BECT ) BECTON DICKINSON CO.
PI   Shank DD,  Spears PA;
DR   WPI; 93-378844/48.
PT   New oligo:nucleotide probes specific for Mycobacteria - used for
PT   detection and amplification of Mycobacteria nucleic acid in
PT   samples
PS   Claim 3; Page 14; 23pp; English.
CC   Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC   (Q51735).  It hybridized to all spp. of mycobacteria tested, but
CC   cross reacted to a few non-mycobacterial spp.  The probe may
CC   be useful as an initial screen for mycobacterial infection.
CC   See also Q51735-45 and Q51747-59.
```

```
SQ   Sequence   91 BP;    5 A;    17 C;    15 G;    4 T;

  Query Match           76.2%;  Score 32;  DB 9;  Length 91;
  Best Local Similarity  0.0%;  Pred. No. 1.03e-06;
  Matches    0;  Conservative   35;  Mismatches   3;  Indels   0;  Gaps   0;

Db     20 hvvshhhsvhhvvhhvhvsvvvvhhvvhvvhhvhyhvy 57
          ::::::: ::::::: :::::::::::::::::: :::
Cp    359 TACCTTCAGAAGATTTACCAGAACCAGAGGTAGAACCT 322


RESULT   11
ID   Q43288 standard; DNA; 744 BP.
AC   Q43288;
DT   27-SEP-1993  (first entry)
DE   Sequence encoding a multivalent antigen-binding single-chain protein
DE   comprising the CC49 VL region connected through the 212 linker
DE   polypeptide to the 4-4-20 VH region.
KW   Antibody; multivalent; variable region; heavy chain; light chain;
KW   linker; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   cds             1..744
FT                   /*tag=  a
FT   cds             1..339
FT                   /*tag=  b
FT                   /product= CC49 VL
FT   cds             340..381
FT                   /*tag=  c
FT                   /product= 212 linker
FT   cds             382..732
FT                   /*tag=  d
FT                   /product= 4-4-20 VH
PN   WO9311161-A.
PD   10-JUN-1993.
PF   20-NOV-1992; U09965.
PR   25-NOV-1991; US-796936.
PA   (ENZO-) ENZON INC.
PI   Bird RE,  Filpula D,  Hardman K,  Rollence M,  Whitlow MD,
PI   Wood JF;
DR   WPI; 93-196999/24.
DR   P-PSDB; R37646.
PT   New multivalent antigen-binding protein e.g. contg. CC49 Mab
PT   regions - useful in diagnosis, for destroying blood clots and
PT   targetting cytotoxic agents or enzymes to tumour cells
PS   Example; Fig 10B; 118pp; English.
CC   "4-4-20 VH" means the variable region of the heavy chain of the
CC   4-4-20 mouse monoclonal antibody (Bird, R.E. et al., Science 242:423
CC   (1988)). The number "212" refers to a specific 14-residue
CC   polypeptide linker that links the 4-4-20 VH and the CC49 VL. See
CC   Bedzyk, W.D. et al., J. Biol. Chem. 265:18615-18620 (1990). "CC49 VL"
CC   is the variable region of the light chain of the CC49 antibody,
CC   which binds to the TAG-72 antigen.
SQ   Sequence   744 BP;    199 A;    164 C;    185 G;    196 T;

  Query Match           66.7%;  Score 28;  DB 7;  Length 744;
  Best Local Similarity 83.3%;  Pred. No. 1.24e-04;
  Matches   35;  Conservative    0;  Mismatches   7;  Indels   0;  Gaps   0;

Db    339 aggctctacttccggtagcggcaaatcttctgaaggtaaagg 380
          ||| ||||| || |||   || ||||||||||||||||||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT   12
ID   Q51746 standard; cDNA; 91 BP.
AC   Q51746;
DT   31-MAY-1994  (first entry)
DE   Oligonucleotide probe MK14-A
KW   Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW   ss.
OS   Synthetic.
PN   EP-571911-A.
PD   01-DEC-1993.
PF   24-MAY-1993; 108325.
PR   26-MAY-1992; US-889651.
PA   (BECT ) BECTON DICKINSON CO.
PI   Shank DD,  Spears PA;
DR   WPI; 93-378844/48.
PT   New oligo:nucleotide probes specific for Mycobacteria - used for
PT   detection and amplification of Mycobacteria nucleic acid in
PT   samples
PS   Claim 3; Page 14; 23pp; English.
CC   Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC   (Q51735).  It hybridized to all spp. of mycobacteria tested, but
CC   cross reacted to a few non-mycobacterial spp.  The probe may
CC   be useful as an initial screen for mycobacterial infection.
CC   See also Q51735-45 and Q51747-59.
SQ   Sequence   91 BP;    5 A;    17 C;    15 G;    4 T;

  Query Match           59.5%;  Score 25;  DB 9;  Length 91;
  Best Local Similarity  0.0%;  Pred. No. 4.10e-03;
  Matches    0;  Conservative   31;  Mismatches   6;  Indels   0;  Gaps   0;

Db     13 vhsyyvvhvvshhhsvhhvvhhvhvsvvvvhhvvhvv 49
          :: :: :::: :::::::::::: :: :::: ::::::
Qy    327 CTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT   13
ID   T29058 standard; DNA; 797 BP.
AC   T29058;
DT   03-OCT-1996  (first entry)
DE   CC49/212 SCA PLAP DNA construct.
KW   Single-chain antigen-binding fusion protein; SCA; immunoeffector;
KW   cytostatic; phospholipase A activating protein; PLAP; cancer;
KW   therapy; monoclonal antibody; CC49; ds.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   cds             1..789
FT                   /*tag=  a
FT                   /product= CC49 VL-212 linker-VH-PLAP fusion
PN   WO9611955-A1.
PD   25-APR-1996.
PF   13-OCT-1995; U12840.
PR   13-OCT-1994; US-323445.
PA   (ENZO-) ENZON INC.
PI   Filpula D,  Shorr R,  Whitlow M;
DR   WPI; 96-221949/22.
DR   P-PSDB; R97380.
PT   New antigen binding proteins contg. immunoeffector or cytolytic
PT   peptide - attached to variable regions of antibody light or heavy
PT   chain, useful in diagnosis and in destroying cancer cells
PS   Example 1; Fig 4A-B; 72pp; English.
CC   A novel DNA construct (T29058) codes for a single-chain antigen-
CC   binding fusion protein (R97380) comprising the light chain variable
CC   region (VL) of monoclonal antibody CC49 linked to the CC49 VH
CC   region via a spacer peptide (R97379). with a C-terminal
CC   phospholipase A activating protein (PLAP) immunoeffector peptide
CC   (R97368).  It was obtd. by ligating a synthetic PLAP fragment to an
CC   existing CC49/212 gene, and can be used to produce the fusion
CC   protein in transformed host cells.  The fusion protein combines
CC   the tumour antigen-binding capability of CC49 with the
CC   immunoeffector function of PLAP.
SQ   Sequence   797 BP;    196 A;    203 C;    195 G;    203 T;

  Query Match           57.1%;  Score 24;  DB 22;  Length 797;
  Best Local Similarity 78.6%;  Pred. No. 1.29e-02;
  Matches   33;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    339 aggctctacttccggtagcggcaaatcctctgaaggcaaagg 380
          ||| ||||| || |||   || ||||| |||||||| |||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT   14
```

```
ID   T29059 standard; DNA; 803 BP.
AC   T29059;
DT   03-OCT-1996  (first entry)
DE   PLAP CC49/212 SCA DNA construct.
KW   Single chain antigen-binding fusion protein; SCA; immunoeffector;
KW   cytostatic; phospholipase A activating protein; PLAP; cancer;
KW   therapy; monoclonal antibody; CC49; ds.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   cds             1..795
FT                   /*tag= a
FT                   /product= PLAP CC49 VL-212 linker-VH fusion
PN   WO9611955-A1.
PD   25-APR-1996.
PF   13-OCT-1995; U12840.
PR   13-OCT-1994; US-323445.
PA   (ENZO-) ENZON INC.
PI   Filpula D,  Shorr R,  Whitlow M;
DR   WPI; 96-221949/22.
DR   P-PSDB; R97381.
PT   New antigen binding proteins contg. immunoeffector or cytolytic
PT   peptide - attached to variable regions of antibody light or heavy
PT   chain, useful in diagnosis and in destroying cancer cells
PS   Disclosure; Page 39-40; 72pp; English.
CC   A novel DNA construct (T29059) codes for a single-chain antigen
CC   binding fusion protein (R97381) comprising the light chain variable
CC   region (VL) of monoclonal antibody CC49 linked to the CC49 VH
CC   region via a spacer peptide (R97379) and with an N-terminal
CC   phospholipase A activating protein (PLAP) immunoeffector peptide
CC   (R97368).  It can be used to produce the fusion protein in
CC   transformed host cells.  The fusion protein combines the tumour
CC   antigen-binding capability of CC49 with the immunoeffector function
CC   of PLAP.
SQ   Sequence   803 BP;    198 A;    205 C;    196 G;    204 T;

  Query Match           57.1%;  Score 24;  DB 22;  Length 803;
  Best Local Similarity  78.6%;  Pred. No. 1.29e-02;
  Matches    33;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    408 aggctctacttccggtagcggcaaatcctctgaaggcaaagg 449
          ||| ||||| || |||   || ||||| |||||||| |||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT   15
ID   Q43291 standard; DNA; 1460 BP.
AC   Q43291;
DT   27-SEP-1993  (first entry)
DE   Sequence encoding divalent CC49/212 single-chain antigen-binding
DE   protein.
KW   Antibody; multivalent; variable region; heavy chain; light chain;
KW   linker; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   cds             1..1458
FT                   /*tag= a
FT   cds             1..339
FT                   /*tag= b
FT                   /product= CC49 VL
FT   cds             340..374
FT                   /*tag= c
FT                   /product= 212 linker
FT   cds             375..426
FT                   /*tag= d
FT                   /product= CC49 VH
FT   cds             427..1065
FT                   /*tag= e
FT                   /product= CC49 VL
FT   cds             1066..1107
FT                   /*tag= f
FT                   /product= 212 linker
FT   cds             1108..1449
FT                   /*tag= g
FT                   /product= CC49 VH
PN   WO9311161-A.
PD   10-JUN-1993.
PF   20-NOV-1992; U09965.
PR   25-NOV-1991; US-796936.
PA   (ENZO-) ENZON INC.
PI   Bird RE,  Filpula D,  Hardman K,  Rollence M,  Whitlow MD,
PI   Wood JF;
DR   WPI; 93-196999/24.
DR   P-PSDB; R37649.
PT   New multivalent antigen-binding protein e.g. contg. CC49 Mab
PT   regions - useful in diagnosis, for destroying blood clots and
PT   targetting cytotoxic agents or enzymes to tumour cells
PS   Example; Fig 16A; 118pp; English.
CC   The number 212 refers to a 14-residue linker. CC49 is the CC49
CC   antibody.
SQ   Sequence   1460 BP;    363 A;    364 C;    365 G;    368 T;

  Query Match           57.1%;  Score 24;  DB 7;  Length 1460;
  Best Local Similarity  78.6%;  Pred. No. 1.29e-02;
  Matches    33;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    339 aggctctacttccggtagcggcaaatcctctgaaggcaaagg 380
          ||| ||||| || |||   || ||||| |||||||| |||||
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


Search completed: Sat Jan  9 17:20:55 1999
Job time : 25 secs.
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Jan 9 17:13:00 1999; MasPar time 105.04 Seconds
716.018 Million cell updates/sec

Tabular output not generated.

Title: >US-08-704-178-1
Description: (322-363) from US08704178.seq (4 of 4)
Perfect Score: 42
N.A. Sequence: 322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363
Comp: TCCAAGATGGAGACCAAGACCATTTAGAAGACTTCCATTTCC

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl-est55
1:em_est1 2:em_gss1 3:em_gss2 4:em_gss3
Database: genbank-est107
5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est2 17:gb_est20
18:gb_est21 19:gb_est3 20:gb_est4 21:gb_est5 22:gb_est6
23:gb_est7 24:gb_est8 25:gb_est9 26:gb_gss1 27:gb_gss2
28:gb_gss3 29:gb_gss4

Statistics: Mean 7.811; Variance 1.736; scale 4.498

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
	1	20	47.6	243	27	FR0033790	Fugu rubripes GSS sequ	6.51e-05
	2	20	47.6	252	12	AA754459	97SN1787 Rice Immature	6.51e-05
	3	20	47.6	477	11	AA708932	z164d01.s1 Soares preg	6.51e-05
	4	19	45.2	217	16	H41340	yn91f01.r1 Homo sapien	1.05e-03
	5	19	45.2	229	21	W35023	mc35f03.r1 Soares mous	1.05e-03
c	6	19	45.2	247	12	AA754458	97SN1784 Rice Immature	1.05e-03
	7	19	45.2	259	7	HUM184D07B	Human fetal brain cDNA	1.05e-03
c	8	19	45.2	337	13	C87905	Mus musculus fertilize	1.05e-03
	9	19	45.2	363	28	AQ134058	HS_3056_A1_B09_MR CIT	1.05e-03
	10	19	45.2	411	26	B46365	HS-1063-B2-B01-MF.abi	1.05e-03
	11	19	45.2	430	16	H10208	ym02d08.s1 Homo sapien	1.05e-03
	12	19	45.2	449	23	AA104834	mp40f09.r1 Barstead MP	1.05e-03
c	13	19	45.2	473	7	AA444530	ve74h09.r1 Soares mous	1.05e-03
c	14	19	45.2	588	10	AA660633	00521 MtRHE Medicago t	1.05e-03
	15	19	45.2	619	27	FR0033832	Fugu rubripes GSS sequ	1.05e-03
	16	19	45.2	718	27	AG011059	Homo sapiens genomic D	1.05e-03
c	17	18	42.9	202	27	AQ062251	CIT-HSP-2347N20.TF CIT	1.54e-02
c	18	18	42.9	263	23	AA208125	mv82b05.r1 GuayWoodfor	1.54e-02
c	19	18	42.9	288	13	AA825949	oe23f10.s1 NCI_CGAP_Br	1.54e-02
c	20	18	42.9	339	8	AA243879	zr65e02.r1 Soares NhHM	1.54e-02
c	21	18	42.9	357	10	AA294453	SWOv3MCA654SK Onchocer	1.54e-02
	22	18	42.9	358	22	AA118186	MBAFCE0B07T3 Brugia ma	1.54e-02
c	23	18	42.9	375	16	H41589	yn91f01.s1 Homo sapien	1.54e-02
	24	18	42.9	376	16	R68647	yi14a03.s1 Homo sapien	1.54e-02
c	25	18	42.9	387	29	AQ245660	HS_2064_A1_E09_T7 CIT	1.54e-02
	26	18	42.9	411	17	AI069472	BSBmMFSZ05N1SK Brugia	1.54e-02
c	27	18	42.9	440	5	T68623	yc43a10.r1 Homo sapien	1.54e-02
	28	18	42.9	450	6	AA406930	MBAFCZ7A01T3 Brugia ma	1.54e-02
c	29	18	42.9	450	15	AA976072	on32g09.s1 NCI_CGAP_Lu	1.54e-02
	30	18	42.9	451	18	AI156727	ue53h05.r1 Soares mous	1.54e-02
c	31	18	42.9	459	18	AI082365	oy67h06.x1 NCI_CGAP_CL	1.54e-02
c	32	18	42.9	462	18	AI082370	oy67h12.x1 NCI_CGAP_CL	1.54e-02
c	33	18	42.9	479	28	AQ121001	HS_3073_A2_B05_MF CIT	1.54e-02
	34	18	42.9	481	13	AA841018	MB3D6AA2E10T3 Brugia m	1.54e-02
	35	18	42.9	494	18	AI183149	uc48g10.r1 Soares 2NbM	1.54e-02
	36	18	42.9	497	17	AI130764	qc12f02.x1 Soares_feta	1.54e-02
	37	18	42.9	509	10	AA416153	vf37a03.r1 Soares mous	1.54e-02
	38	18	42.9	510	13	AA630237	ab96f08.s1 Stratagene	1.54e-02
	39	18	42.9	521	7	AA462447	vg67f07.r1 Soares mous	1.54e-02
	40	18	42.9	526	13	AA842376	MBAFCW7F03T3 Brugia ma	1.54e-02
	41	18	42.9	595	28	AQ159329	nbxb0013L05r CUGI Rice	1.54e-02
	42	18	42.9	636	28	AQ161005	nbxb0006G12f CUGI Rice	1.54e-02
c	43	18	42.9	652	15	AI066075	BSBmMFSZ01L5SK Brugia	1.54e-02
	44	18	42.9	660	11	AF034177	Homo sapiens ntcon6 co	1.54e-02
	45	18	42.9	749	15	AI066008	BSBmMFSZ07D5SK Brugia	1.54e-02

ALIGNMENTS

RESULT 1
LOCUS FR0033790 243 bp DNA GSS 27-JUN-1998
DEFINITION Fugu rubripes GSS sequence, clone 184G17aD8, genomic survey sequence.
ACCESSION AL030157
NID g3272271
KEYWORDS GSS; genome survey sequence.
SOURCE Fugu rubripes.
ORGANISM Fugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontoidei; Tetraodontidae; Fugu.
REFERENCE 1 (bases 1 to 243)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrania,Y., Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-1998)·MRC Human Genome Mapping Project Resource Centre, Hinxton, Cambridge, CB10 1SB, UK. Email: biohelp@hgmp.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid
PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic sequence.
FEATURES Location/Qualifiers
source 1. .243
/organism="Fugu rubripes"
/db_xref="taxon:31033"
/clone_lib="cosmid 184G17"
/clone="184G17aD8"
BASE COUNT 38 a 71 c 33 g 93 t 8 others
ORIGIN

Query Match 47.6%; Score 20; DB 27; Length 243;
Best Local Similarity 85.2%; Pred. No. 6.51e-05;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Db    206 TTCTTCCTCTGGTTCTGANACATCTTC 232
          |||| |||||||||||  | ||||||
Qy    325 TTCTACCTCTGGTTCTGGTAAATCTTC 351


RESULT    2
LOCUS       AA754459      252 bp    mRNA            EST       20-JAN-1998
DEFINITION  97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
            cDNA clone 97SN1787, mRNA sequence.
ACCESSION   AA754459
NID         g2801165
KEYWORDS    EST.
SOURCE      rice.
  ORGANISM  Oryza sativa
            Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
            Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
            Poales; Poaceae; Oryza.
REFERENCE   1  (bases 1 to 252)
  AUTHORS   Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R, Moon,E.P.,
            Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
            Lee,M.C. and Eun,M.Y.
  TITLE     Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
  JOURNAL   Unpublished (1998)
COMMENT
            Contact: Eun M.Y.
            Department of Cytogenetics
            National Inst. of Agri. Sci. and Tech, RDA
            Suwon, Kyunggido, Korea
            Tel: 82 331 290 0301
            Fax: 82 331 290 0307
            Email: myeun@sun20.asti.re.kr
            Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
            University, Yongin, Korea. 449-728 bhnahm@bioserver.myongji.ac.kr
            Seq primer: M13 Reverse Primer.
FEATURES              Location/Qualifiers
     source           1. .252
                      /organism="Oryza sativa"
                      /cultivar="Milyang23"
                      /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
                      XhoI; Directional cDNA library inserted into lambda ZAPII
                      vector  at 5'end with EcoRI and 3' end with Xho I site."
                      /db_xref="taxon:4530"
                      /clone="97SN1787"
                      /clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
                      /tissue_type="Immature Seed"
                      /dev_stage="5 days after pollination"
                      /lab_host="E. coli SOLR"
BASE COUNT        5 a     21 c     12 g     35 t    179 others
ORIGIN

  Query Match           47.6%;  Score 20;  DB 12;  Length 252;
  Best Local Similarity   6.2%;  Pred. No. 6.51e-05;
  Matches     2;  Conservative    22;  Mismatches   8;  Indels   0;  Gaps   0;

Db     99 RBVTRMAHYHDYTNCBBYNNNDYHMWHBBMYB 130
          :::| : :::::|  :::   : :::::::::
Qy    322 AGGTTCTACCTCTGGTTCTGGTAAATCTTCTG 353


RESULT    3
LOCUS       AA708932      477 bp    mRNA            EST       24-DEC-1997
DEFINITION  z164d01.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone
            506689 3', mRNA sequence.
ACCESSION   AA708932
NID         g2718850
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 477)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES              Location/Qualifiers
     source           1. .477
                      /organism="Homo sapiens"
                      /note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I;
                      Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                      oligo(dT) primer [5'
                      AACTGGAAGAATTCGCGGCCGCCTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. Library
                      went through one round of normalization.  Library
                      constructed by M. Fatima Bonaldo."
                      /db_xref="GDB:3812708"
                      /db_xref="taxon:9606"
                      /clone="506689"
                      /clone_lib="Soares pregnant uterus NbHPU"
                      /sex="female"
                      /dev_stage="adult"
                      /lab_host="DH10B"
BASE COUNT      144 a     88 c     96 g    149 t
ORIGIN

  Query Match           47.6%;  Score 20;  DB 11;  Length 477;
  Best Local Similarity  88.5%;  Pred. No. 6.51e-05;
  Matches    23;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db    353 TCTACCTCTGGTTCTGGCTTATCTTC 378
          ||||||||||||||||   ||||||
Qy    326 TCTACCTCTGGTTCTGGTAAATCTTC 351


RESULT    4
LOCUS       H41340        217 bp    mRNA            EST       31-JUL-1995
DEFINITION  yn91f01.r1 Homo sapiens cDNA clone 175801 5' similar to SP:A45036
            A45036 PUR ALPHA=SINGLE-STRANDED-DNA-BINDING PROTEIN - ;.
ACCESSION   H41340
NID         g917392
KEYWORDS    EST.
SOURCE      Homo sapiens clone=175801 library=Soares adult brain N2b5HB55Y
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsite1=Not I Rsite2=Eco RI
            55-year old male. 1st strand cDNA was primed with a Not I -
            oligo(dT) primer [5'
            TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTT 3'],
            double-stranded cDNA was size selected, ligated to Eco RI adapters
            (Pharmacia), digested with Not I and cloned into the Not I and Eco
            RI sites of a modified pT7T3 vector (Pharmacia). Library went
            through one round of normalization to a Cot = 53. Library
            constructed by Bento Soares and M.Fatima Bonaldo. The adult brain
            RNA was provided by Dr. Donald H. Gilden. Tissue was acquired 17-18
            hours after death which occurred in consequence of a ruptured
            aortic aneurysm. RNA was prepared from a pool of tissues
            representing the following areas of the brain: frontal, parietal,
            temporal and occipital cortex from the left and right hemispheres,
            subcortical white matter, basal ganglia, thalamus, cerebellum,
            midbrain, pons and medulla.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
```

```
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 217)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand.
FEATURES              Location/Qualifiers
     source           1. .217
                      /organism="Homo sapiens"
                      /clone="175801"
BASE COUNT       29 a     65 c     51 g     61 t     11 others
ORIGIN

  Query Match           45.2%;  Score 19;  DB 16;  Length 217;
  Best Local Similarity  80.0%;  Pred. No. 1.05e-03;
  Matches    24;  Conservative    0;  Mismatches   6;  Indels   0;  Gaps   0;

Db    100 GNTTCTCCCTCTGCTTCTCTTGAATCTTCT 129
          | |||| |||||| ||||  | ||||||||
Qy    323 GGTTCTACCTCTGGTTCTGGTAAATCTTCT 352


RESULT    5
LOCUS       W35023         229 bp    mRNA            EST       11-SEP-1996
DEFINITION  mc35f03.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 350525
            5', mRNA sequence.
ACCESSION   W35023
NID         g1316977
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 229)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:222325
            Seq primer: mob.REGA+ET
            High quality sequence stop: 204.
FEATURES              Location/Qualifiers
     source           1. .229
                      /organism="Mus musculus"
                      /note="Vector: pT7T3D (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library went through one round of
                      normalization to a Cot = 5. Library constructed by Bento
                      Soares and M.Fatima Bonaldo. RNA was kindly provided by
                      Dr. Minoru Ko (Wayne State University)."
                      /db_xref="taxon:10090"
                      /clone="350525"
                      /clone_lib="Soares mouse p3NMF19.5"
                      /dev_stage="19.5 dpc total fetus"
                      /lab_host="DH10B (ampicillin resistant)"
     mRNA             <1. .>229
BASE COUNT       62 a     63 c     56 g     48 t
ORIGIN

  Query Match           45.2%;  Score 19;  DB 21;  Length 229;
  Best Local Similarity  77.1%;  Pred. No. 1.05e-03;
  Matches    27;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db    162 CTACCTCGTGTTGTCCTGACTCCTCTGAAGGTAAA 196
          ||||||| ||| |  | | || ||||||||||||
Qy    327 CTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAA 361


RESULT    6
LOCUS       AA754458       247 bp    mRNA            EST       20-JAN-1998
DEFINITION  97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
            cDNA clone 97SN1784, mRNA sequence.
ACCESSION   AA754458
NID         g2801164
KEYWORDS    EST.
SOURCE      rice.
  ORGANISM  Oryza sativa
            Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
            Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
            Poales; Poaceae; Oryza.
REFERENCE   1  (bases 1 to 247)
  AUTHORS   Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R, Moon,E.P.,
            Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
            Lee,M.C. and Eun,M.Y.
  TITLE     Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
  JOURNAL   Unpublished (1998)
COMMENT
            Contact: Eun M.Y.
            Department of Cytogenetics
            National Inst. of Agri. Sci. and Tech, RDA
            Suwon, Kyunggido, Korea
            Tel: 82 331 290 0301
            Fax: 82 331 290 0307
            Email: myeun@sun20.asti.re.kr
            Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
            University, Yongin, Korea. 449-728 bhnahm@bioserver.myongji.ac.kr
            Seq primer: M13 Reverse Primer.
FEATURES              Location/Qualifiers
     source           1. .247
                      /organism="Oryza sativa"
                      /cultivar="Milyang23"
                      /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
                      XhoI; Directional cDNA library inserted into lambda ZAPII
                      vector  at 5'end with EcoRI and 3' end with Xho I site."
                      /db_xref="taxon:4530"
                      /clone="97SN1784"
                      /clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
```

```
                     /tissue_type="Immature Seed"
                     /dev_stage="5 days after pollination"
                     /lab_host="E. coli SOLR"
BASE COUNT        7 a     16 c     21 g     34 t    169 others
ORIGIN

  Query Match           45.2%;  Score 19;  DB 12;  Length 247;
  Best Local Similarity  11.4%;  Pred. No. 1.05e-03;
  Matches     4;  Conservative    22;  Mismatches   9;  Indels   0;  Gaps   0;

Db    113 SVVYVWBTAYCDYBHYBDRANHVDDTRCTNDRGYC 147
          :: :::: |  :::::::::| :::: : | :: :|
Cp    357 CCTTCAGAAGATTTACCAGAACCAGAGGTAGAACC 323


RESULT    7
LOCUS       HUM184D07B    259 bp    mRNA            EST       21-MAY-1996
DEFINITION  Human fetal brain cDNA 5'-end GEN-184D07, mRNA sequence.
ACCESSION   D61368
NID         g963007
KEYWORDS    EST; EST(expressed sequence tag); Human fetal brain; similar to
            none(May 29,1995).
SOURCE      Homo sapiens cDNA to mRNA, clone_lib:Clontech human fetal brain
            polyA+ mRNA (#6535).
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 259)
  AUTHORS   Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M.,
            Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H.,
            Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E., Hirai,Y.,
            Maekawa,H., Shin,S. and Nakamura,Y.
  TITLE     Unpublished(101)
  JOURNAL   Unpublished (1995)
REFERENCE   2  (bases 1 to 259)
  AUTHORS   Fujiwara,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAY-1995) to the DDBJ/EMBL/GenBank databases. Tsutomu
            Fujiwara, Otsuka GEN Research Institute,Otsuka Pharmaceutical
            Co.,Ltd; 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01,
            Japan (Tel:0886-65-2888, Fax:0886-37-1035)
COMMENT     Submitted (30-May-1995) to DDBJ by:
            Tsutomu Fujiwara
            Otsuka GEN Research Institute
            Otsuka Pharmaceutical Co.,Ltd
            463-10 kagasuno Kawauchi-cho
            Tokushima, Tokushima
            771-01
            Japan
            Phone: 0886-65-2888
            Fax  : 0886-37-1035.
FEATURES             Location/Qualifiers
     source          1. .259
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Clontech human fetal brain polyA+ mRNA
                     (#6535)"
BASE COUNT       73 a     35 c     45 g     80 t     26 others
ORIGIN

  Query Match           45.2%;  Score 19;  DB 7;  Length 259;
  Best Local Similarity  59.5%;  Pred. No. 1.05e-03;
  Matches    22;  Conservative    6;  Mismatches   9;  Indels   0;  Gaps   0;

Db    140 CTAACAATKGTYTTGGKWAGTTKGATGATGGTAAAGK 176
          ||| |  |:||: |||::| | :  ||| ||||||||:
Qy    327 CTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAAGG 363


RESULT    8
LOCUS       C87905        337 bp    mRNA            EST       04-MAR-1998
DEFINITION  Mus musculus fertilized egg cDNA 3'-end sequence, clone J0253E04,
            mRNA sequence.
ACCESSION   C87905
NID         g2919862
KEYWORDS    EST; EST(expressed sequence tag).
SOURCE      Mus musculus (strain:C57BL/6J) fertilized one-cell-embryo cDNA to
            mRNA, clone_lib:Mouse fertilized one-cell-embryo cDNA
            clone:J0253E04.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (sites)
  AUTHORS   Ko,M.S.H., Kitchen,J.R., Wang,X., Wang,X., Threat,T.A., Sun,T.,
            DePalma,G.E., Liang,Y., Kargul,G.J., Sharara,R., Paonessa,P.D. and
            Doi,H.
  TITLE     Systematic analyses of genes expressed in fertilized mouse eggs(The
            ERATO/Doi Project at Wayne State University)
  JOURNAL   Unpublished (1998)
REFERENCE   2  (bases 1 to 337)
  AUTHORS   Doi,H. and Ko,M.S.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JAN-1998) to the DDBJ/EMBL/GenBank databases.
            Hirofumi Doi, Doi Bioasymmetry Project, ERATO, Japan Science and
            Technology Corporation (JST); WBG Marive East 12F, 2-6 Nakase,
            Mihama-ku, Chiba-shi, Chiba 261-7112, Japan
            (E-mail:hd@bioa.jst.go.jp, Tel:81-43-299-1351, Fax:81-43-297-7530)
FEATURES             Location/Qualifiers
     source          1. .337
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="J0253E04"
                     /clone_lib="Mouse fertilized one-cell-embryo cDNA"
                     /dev_stage="fertilized one-cell-embryo"
BASE COUNT       67 a     86 c     90 g     93 t      1 others
ORIGIN

  Query Match           45.2%;  Score 19;  DB 13;  Length 337;
  Best Local Similarity  77.1%;  Pred. No. 1.05e-03;
  Matches    27;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db    228 CCCTTACTTTAAAAATATTTCCCAGAACCACGGGT 262
          || |||| || | || |||| |||||||||  |||
Cp    363 CCTTTACCTTCAGAAGATTTACCAGAACCAGAGGT 329


RESULT    9
LOCUS       AQ134058      363 bp    DNA             GSS       22-SEP-1998
DEFINITION  HS_3056_A1_B09_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3056 Col=17 Row=C, genomic survey
            sequence.
ACCESSION   AQ134058
NID         g3525424
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 363)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Furlong,J.,
            Shaker,R., Schmidt,S., Traicoff,R. and Hood,L.E.
  TITLE     Construction of a Characterized Clone Resource for Genomic
            Sequencing
  JOURNAL   Unpublished (1998)
COMMENT
            Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3056  row: C  column: 17
```

```
            Class: BAC ends
            High quality sequence stop: 363.
FEATURES              Location/Qualifiers
     source           1. .363
                      /organism="Homo sapiens"
                      /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                      E-Coli DH10B"
                      /db_xref="taxon:9606"
                      /clone="Plate=3056 Col=17 Row=C"
                      /clone_lib="CIT Approved Human Genomic Sperm Library D"
                      /sex="male"
BASE COUNT      126 a     47 c     68 g    122 t
ORIGIN

  Query Match           45.2%;  Score 19;  DB 28;  Length 363;
  Best Local Similarity  80.6%;  Pred. No. 1.05e-03;
  Matches    25;  Conservative    0;  Mismatches   6;  Indels   0;  Gaps   0;

Db    223 CTACTTGTGGTTCTGTTTGATATTCTGAAGG 253
          |||| | |||||||| |  || |||||||||
Qy    327 CTACCTCTGGTTCTGGTAAATCTTCTGAAGG 357


RESULT   10
LOCUS       B46365         411 bp    DNA             GSS       20-OCT-1997
DEFINITION  HS-1063-B2-B01-MF.abi CIT Human Genomic Sperm Library C Homo
            sapiens genomic clone Plate=CT 796 Col=2 Row=D, genomic survey
            sequence.
ACCESSION   B46365
NID         g2551199
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 411)
  AUTHORS   Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
            Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
  TITLE     Construction of a Characterized Clone Resource for Genomic
            Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
            Tagged Connectors
  JOURNAL   Unpublished (1997)
COMMENT
            Contact: Mahairas GG, Zackrone KD, Hood L
            University of Washington
            Seattle, WA 98195, USA
            Tel: (206) 616-8744
            Fax: (206) 685-7301
            Email: kzackron@u.washington.edu
            Sequence Tagged Connector
            Plate: CT 796  row: D  column: 2
            Class: BAC ends
            High quality sequence stop: 411.
FEATURES              Location/Qualifiers
     source           1. .411
                      /organism="Homo sapiens"
                      /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                      E-Coli DH10B"
                      /db_xref="taxon:9606"
                      /clone="Plate=CT 796 Col=2 Row=D"
                      /clone_lib="CIT Human Genomic Sperm Library C"
                      /sex="M"
BASE COUNT      100 a     91 c     97 g    123 t
ORIGIN

  Query Match           45.2%;  Score 19;  DB 26;  Length 411;
  Best Local Similarity  82.8%;  Pred. No. 1.05e-03;
  Matches    24;  Conservative    0;  Mismatches   5;  Indels   0;  Gaps   0;

Db    316 GTGCTACCTCTGGTCCTTCTAACTCTTCT 344
          || |||||||||||  ||  ||| ||||||
Qy    324 GTTCTACCTCTGGTTCTGGTAAATCTTCT 352


RESULT   11
LOCUS       H10208         430 bp    mRNA            EST       23-JUN-1995
DEFINITION  ym02d08.s1 Homo sapiens cDNA clone 46914 3'.
ACCESSION   H10208
NID         g875030
KEYWORDS    EST.
SOURCE      human clone=46914 library=Soares infant brain 1NIB vector=Lafmid BA
            host=DH10B (ampicillin resistant) primer=Promega -21m13 Rsite1=Not
            I Rsite2=Hind III Whole brain from a 73 days post natal female. 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3']; double-stranded
            cDNA was ligated to Hind III adaptors (Pharmacia), digested with
            Not I and directionally cloned into the Not I and Hind III sites of
            the Lafmid BA vector. Library went through one round of
            normalization. Library constructed by Bento Soares and M.Fatima
            Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 430)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            GDB: G00-419-455
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 360
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES              Location/Qualifiers
     source           1. .430
                      /organism="Homo sapiens"
                      /clone="46914"
BASE COUNT      118 a    101 c     70 g    140 t      1 others
ORIGIN

  Query Match           45.2%;  Score 19;  DB 16;  Length 430;
  Best Local Similarity  77.1%;  Pred. No. 1.05e-03;
  Matches    27;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db    198 TTCAACCCCTGGTTGTGGCACAACTTTTCAAGGTA 232
          ||| ||| |||||| ||| | | ||| | ||||||
Qy    325 TTCTACCTCTGGTTCTGGTAAATCTTCTGAAGGTA 359


RESULT   12
LOCUS       AA104834       449 bp    mRNA            EST       13-FEB-1997
DEFINITION  mp40f09.r1 Barstead MPLRB1 Mus musculus cDNA clone 571721 5', mRNA
            sequence.
ACCESSION   AA104834
NID         g1651018
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 449)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
```

```
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:346369
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 414.
FEATURES              Location/Qualifiers
     source           1. .449
                      /organism="Mus musculus"
                      /strain="BALB/c"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      [CATGGATTCGGTACC], digested with Not I and cloned into the
                      Not I and Eco RI sites of the modified pT7T3 vector.
                      Library constructed by Bob Barstead."
                      /db_xref="taxon:10090"
                      /clone="571721"
                      /clone_lib="Barstead MPLRB1"
                      /sex="mixed"
                      /tissue_type="Kidney"
                      /dev_stage="6 weeks"
                      /lab_host="DH10B"
     mRNA             <1. .>449
BASE COUNT      103 a    134 c    125 g     87 t
ORIGIN

  Query Match           45.2%;  Score 19;  DB 23;  Length 449;
  Best Local Similarity  77.1%;  Pred. No. 1.05e-03;
  Matches    27;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     48 CTACCTCGTGTTGTCCTGACTCCTCTGAAGGTAAA 82
        ||||||| ||| |  | | || ||||||||||||
Qy    327 CTACCTCTGGTTCTGGTAAATCTTCTGAAGGTAAA 361


RESULT   13
LOCUS       AA444530      473 bp    mRNA            EST       03-JUN-1997
DEFINITION  ve74h09.r1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone
            832001 5' similar to gb:X58380 M.musculus HMGI-C mRNA for a nuclear
            phosphoprotein (MOUSE);, mRNA sequence.
ACCESSION   AA444530
NID         g2157470
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 473)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:492217
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 335.
FEATURES              Location/Qualifiers
     source           1. .473
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTT
                      T 3']; double-stranded cDNA was ligated to Eco RI
                      adaptors (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of the modified pT7T3 vector.
                      RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                      constructed and normalized by Bento Soares and M.Fatima
                      Bonaldo."
                      /db_xref="taxon:10090"
                      /clone="832001"
                      /clone_lib="Soares mouse mammary gland NbMMG"
                      /sex="male"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
     mRNA             <1. .>473
BASE COUNT       55 a    156 c    128 g    134 t
ORIGIN

  Query Match           45.2%;  Score 19;  DB 7;  Length 473;
  Best Local Similarity  78.8%;  Pred. No. 1.05e-03;
  Matches    26;  Conservative    0;  Mismatches   7;  Indels   0;  Gaps   0;

Db      9 CCTTCAGAAGATCCATCTGATGCTGAGGTAGAA 41
          ||||||||||||  | | ||  | |||||||||
Cp    357 CCTTCAGAAGATTTACCAGAACCAGAGGTAGAA 325


RESULT   14
LOCUS       AA660633      588 bp    mRNA            EST       10-NOV-1997
DEFINITION  00521 MtRHE Medicago truncatula cDNA 5' similar to
            dihydroflavonol-4-reductase, mRNA sequence.
ACCESSION   AA660633
NID         g2604677
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
            Embryophyta; vascular plants; seed plants; Magnoliophyta;
            Magnoliopsida; Rutanae; Sapindales; Fabaceae; Papilionoideae;
            Medicago.
REFERENCE   1  (bases 1 to 588)
  AUTHORS   Covitz,P.A., Smith,L.S. and Long,S.R.
  TITLE     Expressed sequence tags from a root hair-enriched Medicago
            truncatula cDNA library
  JOURNAL   Unpublished (1997)
COMMENT
            Contact: Long SR
            Department of Biological Sciences and Howard Hughes Medical
            Institute
            Stanford University
            Gilbert Biology, Stanford, CA 94305-5020, USA
            Tel: 650 723 3232
            Fax: 650 725 8309
            Email: fa.srl@forsythe.stanford.edu
            Seq primer: T3.
```

```
FEATURES             Location/Qualifiers
     source          1. .588
                     /organism="Medicago truncatula"
                     /cultivar="Jemalong"
                     /note="Organ: Root; Vector: pBK-CMV; Site_1: EcoRI;
                     Site_2: XhoI; cDNA was synthesized from a pooled mRNA prep
                     from elongating root  hairs (30% w/w) and 2-3cm root tips
                     (70% w/w).  XhoI-oligo-dT  linker-primer and EcoRI
                     adaptors were used. cDNAs was cloned  unidirectionally
                     into lambda ZAP Express (Stratagene), amplified,  and
                     mass-excised into pBK-CMV vector plasmids.   More
                     information is available at http://bio-SRL8.stanford.edu."
                     /db_xref="taxon:3880"
                     /clone_lib="MtRHE"
                     /tissue_type="Root hairs & tips"
                     /dev_stage="2-3 day old seedlings"
BASE COUNT      172 a    110 c    135 g    168 t      3 others
ORIGIN

  Query Match           45.2%;  Score 19;  DB 10;  Length 588;
  Best Local Similarity  74.4%;  Pred. No. 1.05e-03;
  Matches    29;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db    277 CCTGTACCTTCAACAGTTGTACCTAATCCTGAGGTGGAA 315
          ||| ||||||||  || | ||||  | || ||||| |||
Cp    363 CCTTTACCTTCAGAAGATTTACCAGAACCAGAGGTAGAA 325


RESULT   15
LOCUS       FR0033832     619 bp    DNA             GSS       27-JUN-1998
DEFINITION  Fugu rubripes GSS sequence, clone 184G17bF10, genomic survey
            sequence.
ACCESSION   AL030199
NID         g3272313
KEYWORDS    GSS; genome survey sequence.
SOURCE      Fugu rubripes.
  ORGANISM  Fugu rubripes
            Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
            Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
            Tetraodontiformes; Tetraodontoidei; Tetraodontidae; Fugu.
REFERENCE   1  (bases 1 to 619)
  AUTHORS   Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrania,Y.,
            Williams,G. and Brenner,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUN-1998) MRC Human Genome Mapping Project Resource
            Centre, Hinxton, Cambridge, CB10 1SB, UK. Email:
            biohelp@hgmp.mrc.ac.uk
COMMENT     Vector: pBluescript II KS
            V_type: phagemid
            PRIMER: KS
            DESCR:
            One pass dye-terminator sequencing of cosmid cloned genomic
            sequence.
FEATURES             Location/Qualifiers
     source          1. .619
                     /organism="Fugu rubripes"
                     /db_xref="taxon:31033"
                     /clone_lib="cosmid 184G17"
                     /clone="184G17bF10"
BASE COUNT      120 a    157 c     97 g    213 t     32 others
ORIGIN

  Query Match           45.2%;  Score 19;  DB 27;  Length 619;
  Best Local Similarity  85.2%;  Pred. No. 1.05e-03;
  Matches    23;  Conservative    0;  Mismatches   4;  Indels   0;  Gaps   0;

Db    195 TTCTTCCTCTGGTTCTGAAACATCTTC 221
          |||| ||||||||||||  | ||||||
Qy    325 TTCTACCTCTGGTTCTGGTAAATCTTC 351


Search completed: Sat Jan  9 17:14:51 1999
Job time : 111 secs.
```